

Inefficiency as a result of individual overemployment

An analysis of the theoretical prospects that
individuals work too much

Erling Skancke



A thesis presented for the degree of
Master of Economic Theory and Econometrics

Department of Economics
University of Oslo
May 2014

Inefficiency as a result of individual over-employment

An analysis of the theoretical prospects that individuals work too much

© Erling Skancke

2014

Inefficiency as a result of individual overemployment

An analysis of the theoretical prospects that individuals work too much

Erling Skancke

<http://www.duo.uio.no>

Print: Reprosentralen, Universitetet i Oslo

Abstract

The thesis investigates the extent to which Pareto inefficiency may occur as a result of individuals working more than that which is socially optimal. Many of the most important contributions to the literature are discussed and presented, starting with an analysis of externalities resulting from individuals' concern regarding their social status, which is illustrated by a simple model of conspicuous consumption. Several existing models of imperfect information in which agents work too much are subsequently presented. Examples with either moral hazard or adverse selection, or both, are analysed, including models in which agents have career concerns.

After a detailed discussion of many of the previous contributions to the literature, two alternative models are developed. We first consider a simple model of adverse selection in teams in which agents differ solely by their productivity. The market equilibrium involves equal work input from all agents, which implies that a fraction of the population works too little, whereas the remaining fraction works too much. It is shown that the efficient allocation can be sustained under an appropriate policy intervention.

Next a shirking-model is presented in which employees are induced to exert effort due to the fear of losing their job. Workers' contribution to output is a function of both their effort level *and* their amount of work hours. Corresponding to similar findings in the existing literature, our model involves an inefficiently low level of employment. However, under the assumption of decreasing marginal productivity, we also find that the sub-optimal level of employment may cause firms to require their employees to work longer hours than the social planner's first-best choice. We also state a sufficient condition under which the market determination of individual work hours is always sub-optimally high, even when the information costs are taken into account.

Having found that individual overemployment is very much a possible outcome of the market determination of individual work hours, we also discuss the efficiency merits of restricting hours worked. Even when employees report working more than their desired level at the prevailing wage rate, a policy restricting work hours may push the economy off the locus of Pareto efficient equilibria without increasing worker welfare. The potential ineffectiveness of restricting work hours as a means to boost employment is briefly illustrated using the previously developed shirking-model.

Preface

The work with this thesis has been an illuminating experience into the difficulties posed by theoretical research. Though the majority of my work with this project has been fruitless, my many efforts have not been in vain, but have served as a great inspiration for future research. I would like to thank my supervisor Bård Harstad, who has always encouraged me to do better. His many comments and precise remarks have contributed to greatly improving the end result. I would also like to thank the Centre for Equality, Social Organization and Performance (ESOP) for their financial support and for providing me with an office space.

This thesis concludes five years of study at the Department of Economics, University of Oslo. I would like to give special thanks to statistician Harald Goldstein and mathematician Nils Christian Framstad, without whom the successful completion of my master's degree would not have been possible. They have been a profound influence, and their continual support and alacrity in answering my numerous questions over the past years have tremendously improved my understanding of the many issues I have encountered during my studies. Their efforts will not be forgotten.

Table of Contents

1	Introduction	1
2	Consumers with concerns regarding their social status	6
2.1	Relative position and social status in economics	6
2.2	A simple model of conspicuous consumption	9
3	Unobservability of agents' characteristics and actions	14
3.1	Unobservability of agents' actions: Moral hazard	15
3.2	Heterogeneity in labour supply costs: The rat race	21
3.3	Unobservability of agents' talent and productivity	27
4	Expanding the literature on individual overemployment	31
4.1	A simple model of adverse selection in teams	32
4.2	An extension of the Shapiro-Stiglitz model	36
4.2.1	Unemployment and the inadequacy of the Walrasian model	36
4.2.2	Individual work hours in the context of aggregate unemployment . .	39
5	Maximum hours regulation as policy response	49
5.1	Mismatch between the equilibrium allocation and the labour supply curve .	49
5.2	Work-sharing as policy instrument against overemployment	55
6	Concluding remarks	57
	References	59

1 Introduction

The very notion that individuals may work too much might seem quite strange to some. However, this idea is both regularly reflected in the public debate on work hours and work conditions, and is sometimes unambiguously conveyed through the language. A notable example of the latter may be the use of the Japanese term *karoshi*, which was introduced in the 1970's and literally means "death from overwork" (see International Labor Organization (2013) for a short discussion). One recent example of the former is the media attention covering the report by Amnesty International (2013) on the treatment of migrant workers employed in Qatar's construction sector for the preparations of the 2022 FIFA World Cup. The report presents several detailed accounts of extremely harsh working conditions, including work days which exceed 12 hours, poor access to proper medical care, and inadequate provision of water. Data on hospital admittance due to work-related injuries are also discussed. Several comments suggest that the provision of simple work equipment, such as protective headgear, would substantially decrease the severity of many of the said injuries. The report frequently makes use of terminology such as *abused*, *exploited* and even *forced* labour.

Despite the fact that the use of notions such as *overwork* frequently appears in the public debate these terms do not seem to be part of the standard terminology used in economic theory. This could suggest that *overwork* is of relatively minor relevance in economic analysis. And indeed, within the classical framework of economics and under the standard assumptions of the perfect market model, *overwork* is a phenomenon which is unlikely to occur. As Cahuc and Postel-Vinay (2005) put it:

For most economists, it is difficult to imagine that agents could work "too much". If they were, then they would obviously reduce their working hours thereby eliminating this inefficiency through individual choice.

Implicit in the above argument is the idea that individuals find themselves in an economic context where all the conditions of the perfect market are satisfied, much like in the Arrow-Debreu model of perfect competition. However, by relaxing some of the assumptions of the latter model, economists have long since identified situations in which failure by the market to ensure an efficient allocation of resources occurs. The most frequently cited sources of market failure move along three dimensions; (i) the presence of *externa-*

lities, (ii) agents acting under *imperfect information*, and (iii) agents taking advantage of *market power*. In this thesis we will perform efficiency analyses of the market determination of individual work hours and effort levels, and in particular discuss the extent to which the above three factors may lead to inefficient market outcomes where individuals either work too many hours or exert too much effort. The efficiency assessment will be based solely on the Pareto criterion. Accordingly, we will define *individual overemployment* as the occurrence of a market equilibrium which is Pareto dominated by an alternative allocation which is itself Pareto optimal, in which individuals work fewer hours or exert less effort. It should be noted that, even though the focus of this thesis will be on variables representing individuals' work hours and effort levels, these variables could in many circumstances be interpreted as a more general measure of working conditions, of which work hours are a part.

Market equilibria characterised by individual overemployment as we have defined it resulting, from either of the above three sources of potential market failure, have already been described by researchers on several occasions. In this thesis we will present and discuss some of the most central contributions to this literature. In our discussion of externalities we restrict attention to the case where individuals' preferences do not only depend on their absolute consumption level, but also on their consumption level relative to other members of society. We will follow Ljungqvist and Uhlig (2000) and present a simple static model of conspicuous consumption, an idea originally put forth by Veblen (1899). The model shows that, as a result of individuals attempting to increase their relative consumption levels, financed by increasing their labour supply, the economy reaches an equilibrium in which all individuals could be better off if everyone both consumed and worked less. The inefficiency results from the fact that individuals do not take into account the negative effect of an increase in their personal consumption level on other individuals: By increasing one's own consumption one also reduces the consumption level of others relative to oneself, hence the externality. Efficiency can, however, easily be restored by an appropriate tax on labour income, decreasing individuals' supply of labour.

In the case of imperfect information we will discuss a multitude of different models from the literature: We note that a recurrent assumption in the literature on moral hazard is that the principal and the agent are both risk-neutral. This assumption is often made in order to simplify the calculations. However, the assumption also entails important con-

sequences for the model's conclusions. In particular, over-provision of effort under moral hazard does not occur. Following Laffont and Martimort (2002) we provide an example in which the agent is risk-averse, and the principal prefers to induce a sub-optimally high level of effort under moral hazard, thus showing that over-provision of effort in these models is indeed possible.

The second source of imperfect information that we will consider is due to heterogeneity amongst economic agents which cannot be directly observed by all parties. We present models in which agents may differ along two dimensions; their preferences for leisure, and their productivity. We briefly present a simple model due to Akerlof (1976) which combines both of these dimensions. He considers a model in which agents work in teams. Employers are unable to distinguish between workers and in equilibrium all employees receive the same compensation for their work, which amounts to the average product of the team. Akerlof establishes the existence of an equilibrium in which agents prefer to work beyond the socially desirable level, in order to benefit from the higher productivity of other workers. In equilibrium, however, no workers are able to reap such benefits and the equilibrium is clearly inefficient.

Further examples of individual overemployment occurring as a result of adverse selection are presented, starting with the model by Landers et al. (1996) who also assume that agents differ in their preferences for leisure. They consider the process of promotion from associate to partner in law firms. With the current partners wanting to maximise the value of the firm, they prefer to sell the partnership to associates who will work longer hours. To avoid the partnership being sold to associates with stronger preferences for leisure, the law firm may require their associates to work inefficiently long hours early in their career so as to screen out the attorneys who prefer shorter hours.

Rebitzer and Taylor (1995) consider a shirking-model in which agents' preferences for leisure in addition to their input of effort in production are unobservable. We present some of their main results. In particular, they establish an equilibrium in which firms are unwilling to unilaterally allow workers to choose between contracts stipulating respectively short and long work hours, even though both firms and workers would benefit from a situation where all firms offered such contracts. They conclude that it is possible that the labour market does not provide enough short-hour jobs.

The discussion of the previous models of individual overemployment is concluded with a

presentation of a variation of the *career concern* model by Holmström (1982). He argues that when employers are unable to perfectly observe the true productivity of workers, the latter may have an incentive to increase their labour input if this contributes to the employer estimating their productivity to be higher, in which case workers may expect higher compensation in the future. It may very well be the case that the workers' labour input is inefficiently high early in their career.

Having presented many of the most important contributions to the literature illustrating the occurrence of overemployment, we proceed to develop two alternative examples. Being unsatisfied with the choice of production functions used in some of the above references, we first put up a simple model of adverse selection in teams. We consider a model where a fraction of the population has higher productivity than others. Firms hire agents to work in pairs and since the workers' productivity cannot be observed by employers, the market equilibrium causes high-productivity workers to work too little, whereas low-productivity workers work too much. It is shown that the efficient allocation can be sustained under an appropriate policy intervention.

Our second model is a shirking-model based on Shapiro and Stiglitz (1984) in which unemployment is required to ensure that workers exert effort. They consider the case where workers' contribution to output is a single binary variable indicating their effort level, which cannot be perfectly observed by employers. We include the extra dimension of individual work hours, much like what was done in Rebitzer and Taylor (1995). We confirm the finding that the market equilibrium yields an inefficiently high unemployment rate. However, under the assumption of decreasing marginal productivity we also find that the sub-optimal level of employment may cause the market outcome to stipulate an inefficiently high level of individual work hours as compared to the first-best. We also state a sufficient condition under which the market choice of individual work hours is always sub-optimally high, even when the information costs are taken into account.

We conclude our analysis by discussing the effects of a policy instrument which is sometimes called upon in the public debate, namely the restriction of work hours through maximum hours legislation. This is done in the context of empirical evidence that employees report working longer hours than they prefer at the prevailing hourly wage rate, as presented in Stewart and Swaffield (1997). We argue that such reports should not be taken as evidence of a market inefficiency due to individual overemployment. Furthermo-

re, Naylor (2003) discusses this evidence and finds that a policy intervention consisting solely of a restriction of work hours is unlikely to increase worker welfare. We add to her analysis by noting that, even in the presence of such evidence, a restriction of work hours may very well push the market outcome off the locus of Pareto efficient allocations. As a further exemplification of the potential drawbacks of restricting work hours, which has also been cited as a possible means to boost employment, we illustrate, using our previously developed shirking-model, that increased employment cannot be ensured by the restriction of individual work hours.

Having analysed a range of different models indicating that individual overemployment is a perfectly possible outcome of the market determination of work hours, or agents' choice of effort level, we conclude that overwork as we have defined it is highly relevant to economic analysis. We do, however, note that many of the models which have been discussed better fit the context of jobs found in relatively high-paying sectors. This challenges the view that the asserted exploitation of relatively low-skilled labour as cited above constitutes an economic inefficiency in the Pareto sense. Though it is possible that an alternative set of modelling tools would yield other conclusions in this regard, it may very well be the case that a different ethical framework than the one assumed by the Pareto criterion is better suited to analyse such situations.

The thesis is organised as follows: Section 2 discusses various contributions to economics emphasising the role of consumers' concern for their social status, and presents a simple model of conspicuous consumption. Section 3 presents and discusses some central examples from the literature, in which the context of imperfect information may be the source of individual overemployment. As we note, both models with moral hazard and adverse selection, as well as models with a mixture of the two, may lead to equilibria in which individuals' labour input is inefficiently high. In Section 4 we first develop a simple model of adverse selection in teams. Thereafter we discuss some of the treatment of unemployment in the literature, before a shirking-model is developed which exhibits the potential for inefficiency due to both sub-optimally high unemployment *and* excessively long individual work hours. Section 5 discusses some of the efficiency merits of introducing maximum hours legislation, as well as its potential ineffectiveness in combatting unemployment. Section 6 provides various concluding remarks.

2 Consumers with concerns regarding their social status

The prospects of market failure in the presence of externalities are well-established in the literature. For instance, it is well-known that if production is a source of pollution then the market outcome will often yield an outcome where production is inefficiently high. Since the level of output is generally positively related to the amount of inputs, and in particular the amount of labour used in the production process, it could be argued that the market outcome also exhibits overemployment of labour: Pareto improvements could be achieved by reducing individuals' labour input.

Although this conclusion may be correct, the overemployment in the above example would often be considered an effect secondary to the pollution effect. In this section, therefore, we present and discuss a framework within which individual overemployment represents the sole economic inefficiency resulting from the presence of an externality in the economy. In particular, we will show how the case where consumers care about their relative position in society may lead to situations where individuals' work effort is too high.

2.1 Relative position and social status in economics

The idea that a person's relative wealth may have important consequences for her utility and happiness is by no means new, and goes at least as far back as Karl Marx, who noted that

A house may be large or small; as long as the surrounding houses are equally small it satisfies the social demands for a dwelling. But if a palace rises beside the little house, the little house shrinks into a hut. (As quoted in Easterlin (1974))

This idea was further explored by economists trying to explain the empirical patterns of an economy's aggregate saving rate. It had already been hypothesised that individuals would be inclined to save a larger portion of their income, the larger their income. This was also confirmed in the empirical findings. However, as the country's average income increased, the data did not indicate that the aggregate saving rate increased, as would be expected if the individual saving rate were a function of absolute income. Duesenberry

(1949) proposed the relative income hypothesis as a possible explanation of these results: If a household's saving decision depends positively on the household's income relative to the national average, this would explain the cross-sectional findings that the rich save a larger portion of their income than do the poor. At the same time, increasing a country's average income would not necessarily change the way in which income is distributed, thus leaving each household's income relative to the national average unchanged. In this case the increase in income would not affect the individual household's saving decision, thus also leaving the aggregate saving rate unchanged.

Easterlin (1974) explores this idea further in the context of an empirical analysis of happiness. Using data from national surveys he finds a discernable positive relationship between an individual's happiness level and his income/wealth: The wealthier a person is, the happier she is. This result is consistently observed across countries. However, when one considers the average of the reported happiness in each country, and relates it to the average income in that country, the positive relationship observed at the national level disappears: Whether a country is rich or poor, it does not seem to affect the average happiness of its population. This is often referred to as the Easterlin paradox. In an attempt to explain these seemingly contradicting results, Easterlin makes use of Duesenberry's idea and assumes that individual utility is a function of this person's current expenditure relative to that of other people. He notes:

In the simplest case, in which the expenditures of every person are given equal weight, the utility obtained by a given individual depends on the ratio of his expenditure to the national per-capita average. The farther he is above the average, the happier he is; the farther below, the sadder. Moreover, if the frame of reference is always the current national situation, then an increase in the level of income in which all share proportionately would not alter the national level of happiness. A classical example of the fallacy of composition would apply: An increase in the income of any one individual would increase his happiness, but increasing the income of everyone would leave happiness unchanged. Similarly, among countries, a richer country would not necessarily be a happier country.

This provides an empirical rationale for further investigating the effects of allowing agents' preferences to depend on relative variables in the analysis of economic efficiency. See

Clark et al. (2008) for a more recent discussion of the empirical evidence and the use of relative preferences as an explanation of the findings. It is worth noting that, although Easterlin's argument may seem compelling, it has also been disputed in the literature. See Hagerty and Veenhoven (2003), Luttmer (2005), Veenhoven and Hagerty (2006) and Easterlin et al. (2010) for a discussion.

Although the explanations provided by both Duesenberry (1949) and Easterlin (1974) are consistent with the empirical evidence it is by no means obvious that concerns about relative position in society and its effect on consumer behaviour gives rise to an economic inefficiency, and in particular in terms of individual overemployment. However, as we shall see, this may very well be the case. This line of thought follows the ideas developed by Veblen (1899) and his notion of *conspicuous consumption*. In his view, Man has a strong proclivity for assessing his own well-being through his success in the eyes of his fellow men. Life in a society presents itself as a struggle between individuals to improve their standing in the community through the exhibition of wealth through excessive consumption, especially that of luxurious goods. He writes:

In any community where goods are held in severalty it is necessary, in order to his own peace of mind, that an individual should possess as large a portion of goods as others with whom he is accustomed to class himself; and it is extremely gratifying to possess something more than others. [...] So long as the comparison is distinctly unfavourable to himself, the normal, average individual will live in chronic dissatisfaction with his present lot; and when he has reached what may be called the normal pecuniary standard of the community, or of his class in the community, this chronic dissatisfaction will give place to a restless straining to place a wider and ever-widening pecuniary interval between himself and this average standard. The invidious comparison can never become so favourable to the individual making it that he would not gladly rate himself still higher relatively to his competitors in the struggle for pecuniary reputability.

Even though Veblen also asserts that an individual's pecuniary reputability can be improved upon through the public display of one's excessive leisure, he seems to conclude that the process of urbanisation, in particular, favours excessive consumption as a means of improving one's social esteem:

The means of communication and the mobility of the population now expose the individual to the observation of many persons who have no other means of judging of his reputability than the display of goods (and perhaps of breeding) which he is able to make while he is under their direct observation. [...] In order to impress these transient observers, and to retain one's self-complacency under their observation, the signature of one's pecuniary strength should be written in characters which he who runs may read. It is evident, therefore, that the present trend of the development is in the direction of heightening the utility of conspicuous consumption as compared with leisure.

Although Veblen does not refer to the resulting situation as economically inefficient in the Pareto sense he does characterise the situation as being "wasteful". It is also immediately obvious that for most people this "waste" must be financed through paid work, and it may easily be the case that individuals end up working beyond any socially optimal level. The next subsection formalises this idea. More recent contributions to the literature confirms the potential for economic inefficiency when individuals care about their social status: Frank (1991), Schor (1993), Fisher and Hof (2000), Ljungqvist and Uhlig (2000), Cahuc and Postel-Vinay (2005), and Arrow and Dasgupta (2009) all specifically address the issue of overemployment, Boskin and Sheshinski (1978) is an early contribution establishing that optimal tax rates are higher when individuals care about their social status. Robson (1992) discusses agents' attitudes towards risk when concern about their social status enters their utility function and shows how equilibrium may involve too high risk-taking and gambling. Also see Weiss and Fershtman (1998) for a discussion.

2.2 A simple model of conspicuous consumption

We now formally illustrate how a socially inefficient outcome involving individual overemployment may emerge in the case where individuals not only care about their personal absolute consumption, but also about their consumption level relative to that of others. We follow the static model presented in Ljungqvist and Uhlig (2000). Both they and several of the above-mentioned authors also consider dynamic models and find similar results. For our purposes we restrict attention to the simpler case of static equilibrium: Suppose the economy is composed of n identical individuals, who all act as price-takers. They each have a utility function of the form $u(c_i, c_i/\bar{c}, h_i)$, where c_i denotes individual

i 's consumption of the single good produced in the economy, and h_i denotes his labour supply. \bar{c} denotes the average consumption in the economy, that is, $\bar{c} = \frac{1}{n} \sum_{i=1}^n c_i$. The agents' preferences exhibits the standard features of positive marginal utility of consumption and increasing disutility in work hours. Furthermore, an individual's utility is higher the larger is the ratio of his consumption to the economy's average consumption. Letting a subscript $j = 1, 2, 3$ indicate partial derivatives, we therefore have $u_1, u_2 > 0$ and $u_3 < 0$. The utility function is assumed to be concave, satisfying $u_{11}, u_{22}, u_{33} < 0$. To simplify some of the below arguments, we also assume $u_{12}, u_{32}, u_{13} \leq 0$.

Production is assumed to take place among many price-taking firms producing a homogeneous good, which are all represented by a single firm with production function $f(\sum_{i=1}^n h_i)$. We assume that output exhibits positive but decreasing marginal productivity of labour and is twice differentiable, e.g. satisfying $f'(\cdot) > 0$ and $f''(\cdot) < 0$. The consumption good is taken to be the numeraire. We let w denote the real wage. Every individual is assumed to receive an equal share of the aggregate profits in the economy, which they perceive as a lump-sum transfer. Firms are assumed to maximise profits, which are given by $\pi = f(\sum_{i=1}^n h_i) - w \sum_{i=1}^n h_i$. Faced with n different sources of labour input, it is obvious that only the aggregate labour supply is of interest to the firm. The first order condition for the firms' demand for labour input is thus given by

$$f'(\sum_{i=1}^n h_i) = w \quad (1)$$

which simply states that the marginal product of labour equals the wage rate.

Agents maximise their utility subject to their budget constraint, which is given by $c_i = \pi/n + wh_i$. Inserting this expression into the utility function, and taking the consumption level of the other individuals as given, individual i 's utility can now be written as $u(\pi/n + wh_i, [\pi/n + wh_i]/\bar{c}, h_i)$. Maximising this expression with respect to individual labour supply, the first order condition can be written as:

$$-\frac{u_1}{u_3} - \left(\frac{1}{\bar{c}} - \frac{1}{n} \frac{c_i}{\bar{c}^2} \right) \frac{u_2}{u_3} = \frac{1}{w} \quad (2)$$

Taking into account the fact that all individuals are the same and their choice of optimal consumption and labour supply satisfies the same conditions it is easy to see that in this static equilibrium all agents consume the same amount of the produced good and work the same amount of hours. In particular, this implies $c_i = \bar{c} \forall i$. Combining this with the

representative firm's demand for labour the free-market outcome of privately optimising behaviour satisfies the following condition:

$$-\frac{u_1}{u_3} - \frac{n-1}{n\bar{c}} \frac{u_2}{u_3} = \frac{1}{f'} \quad (3)$$

In equilibrium, individuals supply labour to the point where the number of hours needed to produce one more unit of the consumption good equals the number of hours an individual is willing to give up in order to increase their absolute consumption with one unit *plus* the number of hours he is willing to give up in order to marginally increase his consumption relative to the other members of the economy. It is easy to see that the market equilibrium involves a negative externality: Whenever agent i increases his individual consumption he also increases the average consumption in the economy. Given the consumption of all others, the utility of any other agent j now decreases as his private consumption decreases relative to the economy's average. This further pushes him to increase his own labour supply and thereby his private consumption. As in Ljungqvist and Uhlig (2000), we have the following result:

Proposition 1: *The market equilibrium outcome is Pareto inefficient, and involves all individuals consuming too much. This consumption is financed through a sub-optimally high individual labour supply.*

To formally verify that the free-market equilibrium is Pareto inefficient we briefly consider the social planner's problem. Assume that the social planner has a purely utilitarian perspective putting equal weights to all individuals in the economy. He therefore seeks to maximise the sum of all individuals' utility, that is $\sum_{i=1}^n u(c_i, c_i/\bar{c}, h_i)$, subject to the constraint that aggregate consumption cannot exceed the produced output. Obviously, this constraint must be binding. The Lagrangian corresponding to the social planner's problem and the associated first-order conditions can thus be written as:

$$\sum_{i=1}^n u(c_i, c_i/\bar{c}, h_i) - \lambda \left[c_i - f\left(\sum_{i=1}^n h_i\right) \right] \quad (4)$$

and

$$u_1(c_i, c_i/\bar{c}, h_i) + \frac{1}{\bar{c}} u_2(c_i, c_i/\bar{c}, h_i) - \frac{c_i}{n\bar{c}} \sum_{i=1}^n u_2 = \lambda \quad (5a)$$

$$-u_3(c_i, c_i/\bar{c}, h_i) = \lambda f' \quad (5b)$$

where the two last conditions hold for all i . Combining the two conditions it is clear that the social optimum must satisfy, for any individual i ,

$$-\frac{u_1}{u_3} - \frac{1}{\bar{c}} \frac{u_2}{u_3} + \frac{c_i}{n\bar{c}} \sum_{i=1}^n \frac{u_2}{u_3} = \frac{1}{f'} \quad (6)$$

Due to the decreasing marginal utility of consumption as well as the increasing marginal disutility of labour input it is clear that the social optimum must be characterised by all agents consuming and working the same amount. As a consequence, $c_i = \bar{c}$, $h_i = h$ and $u_j(c_i, c_i/\bar{c}, h_i) = u_j \forall i, j$. Inserting this into the above equation it follows that the condition for social optimum is given by

$$-\frac{u_1}{u_3} = \frac{1}{f'} \quad (7)$$

that is, in optimum each individual should supply labour to the point where the hours increment necessary to produce one additional unit of the consumption good only equals the number of hours the individual is willing to give up to increase his personal consumption of the good by one unit. Let c^* , h^* and c_{opt} , h_{opt} denote the individual levels of consumption and labour supply in free-market equilibrium and the social optimum respectively. Then since $-\frac{u_1}{u_3} < -\frac{u_1}{u_3} - \frac{u_2}{u_3}$ for all points satisfying $c = f(nh)/n$, $-u_1/u_3$ and $-u_2/u_3$ are decreasing and $1/f'(nh)$ is increasing in h along these points,¹ it follows that both $h_{opt} < h^*$ and $c_{opt} < c^*$.² Hence in the market equilibrium all agents are both working and consuming in excess of the social optimum.

Given that the market outcome is socially inefficient one might be interested in knowing whether a policy intervention would be able to restore efficiency. Ljungqvist and Uhlig (2000) show that this indeed is the case. We have the following result which we give without proof:

Proposition 2: *The socially efficient outcome can be attained through a tax on labour-income, τ , given by*

$$\tau = \frac{\frac{n-1}{n\bar{c}} u_2}{u_1 + \frac{n-1}{n\bar{c}} u_2} \in (0, 1), \quad (8)$$

combined with a lump-sum transfer to all individuals.

¹For $j = 1, 2$ we have $\left(-\frac{u_j(f(nh)/n, 1, h)}{u_3(f(nh)/n, 1, h)} \right)' = -\frac{u_{j1}f' + u_{j3}}{u_3} + \frac{u_j}{u_3} \frac{u_{j3}f' + u_{33}}{u_3} < 0$.

²Assume that $h_{opt} \geq h^*$ then $\frac{1}{f'(nh_{opt})} = -\frac{u_1}{u_3}(h_{opt}) < -\frac{u_1}{u_3}(h_{opt}) - \frac{u_2}{u_3}(h_{opt}) \leq -\frac{u_1}{u_3}(h^*) - \frac{u_2}{u_3}(h^*) = \frac{1}{f'(nh^*)}$, contradicting that $1/f'(nh)$ is increasing.

		Jones	
		Few hours, low pay	Long hours, high pay
Smith	Few hours, low pay	Second best for both	Best for Jones, worst for Smith
	Long hours high pay	Best for Smith worst for Jones	Third best for both

Figure 1: Labour supply as a Prisoner's dilemma-type game

The above inefficiency can also easily be illustrated using one of the most well-known tools from non-cooperative game theory, namely that of the Prisoner's dilemma. Specifically, following Frank (1991) we could consider a situation with two different families; the Smiths and the Joneses. Both families care about their consumption level relative to the other. The head of each family faces two choices regarding their labour supply. Either they can choose a job with relatively few work hours but also a quite low pay, or they may choose to work long hours in an other job with correspondingly higher pay. Both families would find it better if they were the only family taking the job with long hours, and both families would find it to be the worst outcome to be alone to choose the job with short hours. However, both families find it second best if they both choose the job with fewer hours. The case where both families choose the long hours job is perceived as the third best outcome for both families. The situation is illustrated in Figure 1.

Now it is easily verified that choosing the job involving long hours, high pay, and therefore also high consumption is a dominant strategy for both families. Consequently, the situation where both families choose this job is the single Nash equilibrium of this game. However, the equilibrium is obviously inefficient, as both families would be better off if they both worked less. By *keeping up with the Joneses* (and the Smiths) the privately

optimal choices of the two families moves the economy to an equilibrium where everyone could be better off if everyone chose to work less. This is an obvious case of an inefficiency caused by individual overemployment. It should be noted that, as in the above model, the inefficiency can easily be eliminated by an appropriate tax on either labour demand and/or supply.

This section has discussed the relevance in the field of economics of the contention that individuals do care about their relative position in society. We noted that some of the early contributions to the literature on the topic made the claim that such a theoretical prediction would be consistent with certain features observed in the economy, thus asserting the empirical relevance of this assumption. We next showed that the inclusion of this assumption in the analysis easily can give way to the potential for a negative external effect on other individuals when agents choose their optimal level of labour supply and consumption. This externality indeed leads to the manifestation of a market inefficiency due to individual overemployment.

3 Unobservability of agents' characteristics and actions

We have already seen that the presence of externalities in economic models can cause the market equilibrium outcome to feature individual overemployment. In this section we investigate the consequences of an other classical source of market failure, namely imperfect information. We present and discuss some of the most central contributions to the literature where imperfect information causes individuals to work beyond the social optimum. We begin our discussion by looking at a model of pure moral hazard based on the principal-agent framework. We show that in the case where the the agent is risk-averse, it may very well be such that his equilibrium effort level exceeds the first-best optimum when effort is not observable.

We next examine previous models in which employers are unable to directly distinguish the personal characteristics of (potential) employees. The subsequent models thus exhibit the potential for adverse selection. We will consider two different dimensions along which agents may differ. First, we will examine the case where agents put different value on their spare time, or equivalently, supplying labour is more costly for some agents than for

others. The second form of heterogeneity stems from agents' differing *talent*, that is, their contribution to the production process for a given level of effort. As we shall see, both these forms of heterogeneity could potentially lead to individuals working too long hours. It is worth noting that these models generally require at least two critical features: Either the models are dynamic in nature, meaning that the time dimension is included, and also exhibit imperfect observability of one of the aforementioned heterogeneities among agents, or the model is static, in which case one needs unobservability along at least two dimensions.

3.1 Unobservability of agents' actions: Moral hazard

The development of the economics of asymmetric information greatly contributed to our understanding of the potential for market failure, and has found its applications in most fields of economics. For instance, in the case of moral hazard, that is, when economic agents' actions are unobservable to other parties, early contributions provided an economic rationale for the absence of insurance markets even when such insurance would be socially beneficial. The main point is that as long as the insurer is risk-neutral and the insuree is risk-averse, the social optimum will always be characterised by full insurance. However, if the uncertainty in the economic environment could be influenced by the insuree's actions, which are unobservable to the insurer, the former might not have an incentive to perform the "right" action. If this is the case, the insurer will not be willing to provide full insurance. See for instance Pauly (1968) for a comment regarding the case of health insurance.

Central to the analysis of moral hazard is thus the determination of the agent's choice of effort, as this was the source of the principal's reluctance to provide full insurance. However, in the case where the agent is risk-averse researchers were only able to establish a few general results and the analysis often proved to be relatively complicated. Many researchers thus chose to confine their attention to the much simpler case in which the agent is also risk-neutral. A result which seems to hold under quite general conditions is that the agent is likely to exert a sub-optimally low effort level under moral hazard, at least as long as the principal is constrained from punishing the agent too hard in the event where a desirable outcome is not achieved. Such a constraint is referred to as limited liability. Moreover, under the assumption that the agent is risk-neutral, insurance is not

an issue, and thus the sole cause of economic inefficiency is the agent's under-provision of effort.

The use of the moral-hazard model with the feature that the principal and the agent are both risk-neutral may seem somewhat misplaced in certain applications. It nevertheless remains a popular modelling choice due to the fact that it greatly simplifies many of the calculations. Banarjee et al. (2002), for instance, use a simple model of the contracting between landlords and tenants to establish the possibility that agricultural production in developing countries may be sub-optimally low due to the tenants' low production input. A similar conclusion is reached in Bardhan and Udry (1999) in their treatment of rural land markets. These analyses thus provide a theoretical justification for a government intervention aimed at increasing agricultural production. One possible intervention would be land reform programs in which land rights are transferred from the landlords to their tenants. The above analyses would in this case predict an increase in agricultural productivity and are therefore sometimes used as the normative basis for such land reforms. Without questioning these reforms in and of themselves, one might still be inclined to doubt the adequacy of a modelling choice in which under-provision of effort is a rather likely outcome, and over-provision of effort a theoretical impossibility.

To illustrate that over-provision of effort is a possible outcome in the theory of moral-hazard we follow a simple model presented in Laffont and Martimort (2002). Consider a contracting game between two individuals, the principal and the agent. The former owns a production technology in which the agent's effort, e , can be used as input. The production outcome is a random variable whose probability distribution is influenced by the agent's effort level. As a simplification, we assume that with probability $\pi(e)$ the production will be high, yielding a value of H , and with probability $1 - \pi(e)$ production will be low and yield a value of $L < H$. Thus, expected output will be given by

$$q = L + \pi(e)(H - L) \quad (9)$$

The principal is assumed to be risk-neutral, whereas the agent is assumed to be risk-averse. Specifically, the agent's preferences are represented by the following utility function

$$u(y, e) = v(y) - c(e), \quad (10)$$

where v is a twice differentiable, strictly increasing and concave functions satisfying $v(0) = 0$. y denotes the transfer received by the agent. We will consider a case where the agent's

effort level is only allowed to take on three different values, that is, $e \in E = \{e_0, e_1, e_2\}$ with $0 = e_0 < e_1 < e_2$. The agent's disutility of effort is represented by the function c , which is assumed to be increasing starting at the origin, i.e. $0 = c(e_0) < c(e_1) < c(e_2)$. Similarly, it is assumed that the probability of a high production outcome depends positively on the effort exerted by the agent and that zero effort always yields the low outcome. We thus have $0 = \pi(e_0) < \pi(e_1) < \pi(e_2)$. The contracting game is assumed to be of the form of a single take it or leave it offer from the principal and we restrict our attention to sub-game perfect Nash equilibria only. The principal thus seeks to find a contract which maximises profits subject to the constraint that the agent accepts the contract. Both parties are assumed to be expected utility maximisers. The outside option of the agent is taken to be zero, meaning that the agent will accept all contracts yielding a non-negative expected utility. This is referred to as the agent's participation constraint. We first consider the case of perfect information where the agent's effort level can be observed. Since the principal is risk-neutral and the agent risk-averse it is easily verified that the principal should bear all the risk, meaning that the agent should receive a certain transfer with probability 1. Keeping this in mind, the principal's maximisation problem under perfect information can be formulated as

$$\max_{e,y} L + \pi(e)(H - L) - y \quad \text{subject to } v(y) - c(e) \geq 0 \quad (11)$$

To make our point in the below discussion on the effect of moral hazard we will assume that the intermediary effort level e_1 is the optimal in the perfect information case. e_1 will thus be the Pareto optimal level of effort. To see when this would be the case, we first note that the participation constraint in the above problem must always be binding in optimum. e_1 is thus the first-best optimal choice of effort if it yields a higher profit to the principal than the alternative effort levels. This will be the case when

$$\begin{aligned} L + \pi(e_1)(H - L) - v^{-1}(c(e_1)) &> L \quad \text{and} \\ L + \pi(e_1)(H - L) - v^{-1}(c(e_1)) &> L + \pi(e_2)(H - L) - v^{-1}(c(e_2)), \end{aligned} \quad (12)$$

which, if combined, are equivalent to

$$\frac{v^{-1}(c(e_1))}{\pi(e_1)} < H - L < \frac{v^{-1}(c(e_2)) - v^{-1}(c(e_1))}{\pi(e_2) - \pi(e_1)} \quad (13)$$

We now turn to the case where the agent's effort level can no longer be observed by the principal. As a consequence, it will no longer be possible to write a contract contingent on

the effort level. Rather, the principal will choose a compensation scheme which depends on the production outcome. Specifically we assume that the agent receives a compensation h if production is high, and a transfer l if production is low. Given the values of h and l , the agent will now choose the effort level that maximises his expected utility. This is referred to as the agent's incentive compatibility constraint. It is immediately clear that as long as $h \leq l$ the agent will always choose the lowest effort level, hence e_0 is always implementable. For our purposes an interesting case occurs when it is no longer possible for the principal to induce the first-best optimal effort level e_1 . As Laffont and Martimort (2002) point out, we have the following result:

Proposition 3: *Assume that (13) holds, and that*

$$\frac{c(e_2)}{c(e_1)} < \frac{\pi(e_2)}{\pi(e_1)} \quad (14)$$

Then e_1 is no longer implementable under moral hazard and with the agent being risk-averse there exists parameter values $H, L, \pi(e_1), \pi(e_2), c(e_1), c(e_2)$ such that the principal prefers to induce the effort level e_2 under imperfect information with limited liability. This means that in equilibrium the agent exerts a sub-optimally high level of effort.

To demonstrate that the above result holds, we first find the conditions under which e_1 would be implementable. Suppose h and l are such that e_1 is the agent's optimal choice of effort, then we must have

$$\begin{aligned} v(l) + \pi(e_1)[v(h) - v(l)] - c(e_1) &\geq v(l) \quad \text{and} \\ v(l) + \pi(e_1)[v(h) - v(l)] - c(e_1) &\geq v(l) + \pi(e_2)[v(h) - v(l)] - c(e_2), \end{aligned} \quad (15)$$

which, if combined, is equivalent to

$$\frac{c(e_1)}{\pi(e_1)} \leq v(h) - v(l) \leq \frac{c(e_2) - c(e_1)}{\pi(e_2) - \pi(e_1)} \quad (16)$$

This means that as long as

$$\frac{c(e_2) - c(e_1)}{\pi(e_2) - \pi(e_1)} < \frac{c(e_1)}{\pi(e_1)}, \quad (17)$$

there exists no pair (h, l) satisfying the agent's incentive compatible constraint for the choice of e_1 , and hence the latter effort level cannot be implemented. However, the above condition is simply equivalent to (14). Assuming the latter equation holds we can derive the conditions under which the principal would prefer to induce the effort level e_2 over e_0 . The lowest effort level e_0 can simply be induced by choosing a compensation scheme

where $h = l = 0$, in which case the principal earns a profit equal to L . Suppose rather that h and l are such that e_2 is the agent's optimal choice of effort. Then we must have

$$\begin{aligned} v(l) + \pi(e_2)[v(h) - v(l)] - c(e_2) &\geq v(l) \quad \text{and} \\ v(l) + \pi(e_2)[v(h) - v(l)] - c(e_2) &\geq v(l) + \pi(e_1)[v(h) - v(l)] - c(e_1), \end{aligned} \quad (18)$$

which, if combined, using (14), can be rewritten as

$$v(h) - v(l) \geq \max \left[\frac{c(e_2) - c(e_1)}{\pi(e_2) - \pi(e_1)}, \frac{c(e_2)}{\pi(e_2)} \right] = \frac{c(e_2)}{\pi(e_2)} \quad (19)$$

We consider the case of limited liability, which in particular means that the compensation in the event where production is low is not allowed to be negative, i.e. $l \geq 0$. It can be verified that the principal would choose l and h so as to make both the incentive compatibility and the participation constraints bind. Inserting the binding incentive compatibility constraint into the agent's participation constraint it now follows that the latter is binding with $l = 0$, meaning that the limited liability constraint would also be satisfied. Given $l = 0$ the principal's optimal choice of h is now given by

$$h = v^{-1} \left(\frac{c(e_2)}{\pi(e_2)} \right) \quad (20)$$

It remains to verify that the principal would prefer to induce the effort level e_2 . Using the above results, this would be the case as long as

$$\begin{aligned} L + \pi(e_2)[H - L] - l - \pi(e_2)(h - l) &> L \\ \Leftrightarrow H - L &> v^{-1} \left(\frac{c(e_2)}{\pi(e_2)} \right) \end{aligned} \quad (21)$$

The conditions for over-provision of effort under moral hazard can now be summarized by three constraints on the parameters of the model; (13) (the condition that e_1 be first-best optimal), (14) (the condition that e_1 cannot be implemented under moral hazard), and (21) (the condition that e_2 be the principal's second-best choice under imperfect information). Combining (13) and (14) we find that a necessary condition is

$$\frac{c(e_2)}{c(e_1)} < \frac{\pi(e_2)}{\pi(e_1)} < \frac{v^{-1}(c(e_2))}{v^{-1}(c(e_1))} \quad (22)$$

It is straightforward to verify that these inequalities cannot hold if v is linear (in which case v^{-1} is also linear). To show that the above condition still may hold under our assumptions we make use of the following result:

Lemma 1: *Let $g : [0, \infty) \rightarrow \mathbb{R}$ be twice differentiable, strictly increasing and strictly convex, satisfying $g(0) = 0$. Then*

$$\frac{x}{y} < \frac{g(x)}{g(y)} \quad \forall x > y > 0 \quad (23)$$

Proof: Define $G : [0, \infty) \rightarrow \mathbb{R}$ by $G(y) = g(y) - yg'(y)$ and note that $G(0) = 0$ and $G'(y) = -yg''(y) < 0$, and thus also $G(y) < 0 \forall y > 0$. From the mean value theorem, and using the fact that g' is increasing and $x > y$, it now follows that for some $t \in (y, x)$ we have

$$\begin{aligned} (x - y)g(y) &< y(x - y)g'(y) < y(x - y)g'(t) = y \int_y^x g'(s)ds = y[g(x) - g(y)] \\ &\Leftrightarrow \frac{x}{y} < \frac{g(x)}{g(y)} \end{aligned} \quad (24)$$

□

From our assumption that the agent is strictly risk-averse with the function v satisfying $v(0)$ it follows that the function v^{-1} satisfies the conditions of the above Lemma. We thus have the following result, which needs no further proof:

Corollary 1: *From our assumptions regarding the function v it must be the case that*

$$\frac{c(e_2)}{c(e_1)} < \frac{v^{-1}(c(e_2))}{v^{-1}(c(e_1))} \quad (25)$$

With appropriate values of $\pi(e_1)$ and $\pi(e_2)$ Equation (22) can thus always be satisfied.

It is worth emphasising the role of the functional form of v in the above result. As we already noted, if v were linear, that is, if the agent were risk-neutral, the above conditions could never hold, which is also one of the reasons why over-provision of effort does not occur in this case.

Equations (14) and (21) can be combined to find the last important constraint on the model's parameters. In particular, it must be the case that

$$v^{-1}\left(\frac{c(e_2)}{\pi(e_2)}\right) < \frac{v^{-1}(c(e_2)) - v^{-1}(c(e_1))}{\pi(e_2) - \pi(e_1)} \quad (26)$$

This condition can be reformulated as a lower bound on the probability $\pi(e_1)$ as a function of $\pi(e_2)$:

$$\pi(e_1) > \pi(e_2) - \frac{v^{-1}(c(e_2)) - v^{-1}(c(e_1))}{v^{-1}\left(\frac{c(e_2)}{\pi(e_2)}\right)} \quad (27)$$

We are now ready to prove the last part of the above Proposition:

Proof of Proposition 3: We only need to establish the existence of a set of parameters such that (13), (14) and (22) all hold simultaneously. Letting $\pi(e_1)$ satisfy (27) and choosing $\pi(e_2) = 1$ it is easily verified that (22) will automatically be satisfied, as will (26) from our choice of $\pi(e_1)$. Now there must exist a number $H - L > 0$ such that

$$v^{-1}(c(e_2)) < H - L < \frac{v^{-1}(c(e_2)) - v^{-1}(c(e_1))}{1 - \pi(e_1)}, \quad (28)$$

and thus all the conditions are satisfied. It follows that the principal prefers to induce e_2 under moral hazard and the equilibrium thus exhibits over-provision of effort relative to the first-best outcome with perfect information. \square

The above model illustrates how moral hazard may be the source of two different inefficiencies: First, from the fact that the agent is risk-averse, the allocation is inefficient due to the absence of full insurance. The equilibrium also involves a second kind of inefficiency, namely the over-provision of effort by the agent. As we have already indicated, neither of these inefficiencies can occur in a model where both the agent and the principal are assumed to be risk-neutral. To the extent that the case of risk-aversion is a more realistic assumption, knowledge about the above result may be valuable in many contexts of economic analysis. In the next two subsections we introduce an other dimension of imperfect information, namely the unobservability of agents' characteristics.

3.2 Heterogeneity in labour supply costs: The rat race

We begin our discussion of the adverse selection models yielding overemployment in equilibrium by considering the case where agents' preferences for leisure differ. The earliest example formalising this idea in the literature is due to Akerlof (1976). His original model examines the determination of a variable representing working conditions, but this variable could also easily be interpreted as individual work hours. The model combines the idea that agents differ in their productive capabilities with the notion that agents may differ in their preferences for leisure. We briefly present (a slight variation of) his model and state his main result:

Consider an economy where there are N different grades of workers, and the higher the grade, the higher is the productivity of the worker. Agents work together in teams to produce output, q . Output per worker depends positively on the team's average grade, \bar{n} , and the number of hours, h , the members of the team stay at work. Specifically, output

is assumed to be given by

$$q = \bar{n} + h \quad (29)$$

Agents' preferences are assumed to be separable and linear in their compensation, y . The agents' preferences for leisure depends negatively on their grade, meaning that workers with higher productivity are assumed to be more willing to trade leisure for consumption. Specifically, the utility, u_n , of a worker of grade n , is given by

$$u_n = y - h - 3/8(h - n)^2, \quad (30)$$

where the choice of the number $3/8$, as Akerlof explains, is to ensure that only workers of the same grade come together in teams in equilibrium.

The economy is composed of firms in perfect competition hiring agents to work in teams with a prespecified number of work hours h , which may correspond to any integer. Workers choose for which firm, or in which team, they want to work. The firms are unable to distinguish between the workers, and since the labour market is perfectly competitive the agents will all be paid the average product of their respective team. As Akerlof (1976) shows, we have the following result:

Proposition 4: *The above model has an equilibrium in which workers of grade $n = 1$ work in one-hour teams, and workers of grade $n > 1$ work in teams where the working hours are $n + 1$. This means that all workers of grade $n > 1$ work more than the Pareto efficient level.*

It is easy to see that efficiency requires all agents to work to the point where their marginal disutility of labour equals their marginal productivity. This means that workers of grade n should work n hours. However, if teams were put together according to the requirements of an efficient allocation, under the specific functional forms used to represent both agents' preferences and production, agents of lower grade would have the incentive to join agents of higher grade and work longer hours than what would be the optimal choice for this grade. Despite the increase in disutility due to longer working hours, this change of work place would be privately profitable to the lower grade workers, as they would benefit from a sufficiently large increase in compensation due to the superior productivity of the higher grade workers. Thus the Pareto efficient allocation cannot be an equilibrium. Note however that in the equilibrium described above none of the workers are able to benefit from the superior quality/productivity of the higher grade classes, as in equilibrium all

workers of the same grade work together. Nevertheless, no individual worker has the incentive to work shorter hours, as this would imply being matched with workers of inferior quality, whose lower productivity would entail a more modest compensation. It is individually optimal to stay with the workers of your own grade. The equilibrium is thus obviously Pareto inefficient.

Akerlof himself characterises his model as being "surrealistic". Even though this could be said about many common models used in economics, one might be inclined to agree that the fact that the numerical example presented in the paper hinges on very specific functional forms and parameter values represents somewhat of a drawback. As Landers et al. (1996) put it:

Akerlof's demonstration that over-work equilibria are possible was presented in a self-consciously unrealistic example. This may have created the unfortunate impression that the rat-race equilibrium was an interesting theoretical example of market failure, but one having little connection to the operation of actual labor markets.

Fortunately, other models which may seem more realistic have also been developed. Landers et al. (1996) themselves propose a model aimed at analysing the recruiting of partners by law firms. The idea is that attorneys may provide legal services either in a spot market with a given productivity level, or the attorneys may be paired up in a limited number of partnerships where the productivity is higher than in the spot market and where the two partners equally share the output that they produce. Some attorneys value their spare time more than others and are in principle not willing to work as much as others. As a consequence, due to the equal sharing rule among partners, it is always preferable to be paired up with a partner who has a low disutility of effort, and are willing to work long hours in a partnership. The model is dynamic, and agents are assumed to live (or to be available in the labour market) for two periods. Attorneys work the first period of their career as associates in a law firm which has an established partnership pair. At the end of the first period the partnership is auctioned off to the two highest-bidding associates, while the remaining attorneys are free to continue to work as associates in the second period. It is straightforward to show that it is both Pareto efficient and privately optimal for the current partners to make sure that the new partners will have a low disutility of work hours and therefore put in a lot of effort in the partnership in the second period.

The problem is thus to ensure that the partnership remains in the hands of attorneys who will maximise its value.

In the full information case where agents' preferences for leisure are observable matters are quite simple: In the first period of their career attorneys work either long or short hours according to their actual preferences. The partnership is then sold to a pair of the attorneys working long hours. In the second period these two attorneys continue the partnership and both work long hours. The other attorneys, now second-period senior associates, continue to work the Pareto optimal level in the spot market, and new first-year associates are hired.

When agents' disutility of effort cannot be directly observed matters are more complicated. In particular, given the contracts described in the full information case there exists parameter values such that attorneys who actually prefer shorter hours will pretend to have a lower disutility of effort and therefore work longer hours in the first period, bid for the partnership and be paired with an attorney who works long hours in the second period. The associate pretending to enjoy long hours will, however, not put in as much effort in the second period as his partner and thus both the value of the partnership and the second-period allocation of effort will be sub-optimal. Clearly, the contracts proposed in the full information case cannot prevail in an equilibrium where all agents work optimal hours. As Landers et al. (1996) point out, we have the following result:

Proposition 5: *The solution to the law firm's problem of selling the partnership to associates maximising the firm's value, is to offer wage contracts to its associates specifying work hours which associates with higher preferences for leisure will not accept. A separating equilibrium emerges in which the only potential buyers of the partnership are lawyers with a proclivity for working longer hours. However, the separating contract may involve the first-period associates working beyond the Pareto optimal level.*

If firms offer wage-hours packages such that no associate with actual preferences for shorter hours has the incentive to pretend to enjoy longer hours, then it will be common knowledge that the partnership will only be owned by attorneys with preferences for longer hours, thus ensuring an optimal allocation of effort amongst the partners. However, constructing such separating contracts may require the longer-working associates to work more than what is socially optimal in the first period, so as to screen out the attorneys with preferences for shorter hours. In the process of ensuring that the partnership in the

future will be owned by attorneys committed to working long hours, the potential buyers of the partnership among the first-period associates are required to work excessively long in order to signal their true preferences for working long hours. The resulting equilibrium clearly entails a Pareto inefficiency due to individual overemployment.

In the above framework of attorneys with differing preferences for working long hours one might easily imagine a situation where law firms would only be interested in hiring associates who choose the long-working package in the first period of their career. One might ask whether attorneys with preferences for shorter hours will at all be able to find a job at a law firm. A similar question was discussed in Rebitzer and Taylor (1995). They consider a shirking-model in which production output depends positively on the employees' input of both effort and work hours. Both inputs entail a disutility to the workers, and since the employees' effort level cannot be perfectly observed by the employers, the former have an incentive to only exert low levels of effort. However, workers who choose to shirk will be detected with a positive probability in which case they will be dismissed from the firm, upon which they leave the market. The employers' problem is thus to construct a cost-minimising contract which ensures that the employees are willing to exert effort, that is, such that the fear of being dismissed outweighs the disutility of exerting effort. See Section 4.2 for a further discussion of the shirking-paradigm.

One of the differences in Rebitzer and Taylor (1995) compared to many of the existing shirking-models is that the former introduce the additional dimension of adverse selection in their model. As in Landers et al. (1996), they assume that there exists two different types of individuals, with one having higher (marginal) disutility of work hours than the other. However, the employer is unable to directly distinguish between the two types. This therefore further complicates the employers' optimal contracting problem. Rebitzer and Taylor argue that an employer faces three possibilities: Either he can offer a *screening contract* which entails such long hours that the short-hour workers will refuse it but still ensures that the long-hour workers do not shirk; or he can offer a *separating contract* including two different alternatives which the workers may choose between, such that long-hour workers work long hours and short-hour workers work short hours. These alternatives ensure that both worker-types exert effort. The employers' last possibility is to offer a *pooling contract* under which both types of workers have to work long hours.³

³A pooling contract where all workers work short hours will always be dominated by the above separating contract, and need not be considered.

Under this contract only the long-hour workers will exert effort, whereas the short-hour workers prefer to shirk.

One of the main points set forth by Rebitzer and Taylor is that a firm's cost per unit of labour, as they put it, "depends crucially on the mix of the two types of workers". Let θ and $1 - \theta$ respectively denote the proportion of short-hour and long-hour workers in the population, and let C^s and C^p denote each firm's unit labour cost associated with the separating and the pooling contract, respectively. These costs are functions of the proportion of short-hour workers that the individual firm attracts, which again depends on both the proportion of short-hour workers in the population as well as the contracts offered by other firms. One of the main results in Rebitzer and Taylor (1995) is thus the following:

Proposition 6: *Assume that all firms in the market offer the pooling contract ensuring that only the no-shirking condition of long-hour workers is satisfied. Then there exists values of θ , the proportion of short-hour workers in the economy, for which no individual firm has an incentive to unilaterally switch to the separating contract even though a simultaneous switch by all firms to such a contract would be both profitable to all firms and increase the utility of all workers.*

The intuition behind the above result is that a firm which unilaterally switches to the separating contract risks attracting a disproportionate amount of short-hour workers, for whom the costs of ensuring that no shirking takes place is higher than for the long-hour workers. It may thus be the case that the firm prefers to continue offering the pooling contract so as to avoid attracting too many of the short-hour workers. On the other hand, if all firms offered the separating contract the short-hour workers would be evenly distributed among firms and thus the individual firm's costs associated with short-hour workers would not necessarily be that high. It is thus possible that all firms would benefit from offering the separating contract provided that all other firms also do so. If this were the case the short-hour workers would work fewer hours and simultaneously choose to exert effort.

Despite the fact that, for the relevant values of θ , both workers and firms would prefer the allocation where all firms offer the separating contract over the equilibrium allocation where all firms offer the pooling contract it is not necessarily the case that the former allocation involves a first-best level of individual work hours. Indeed, due to the no-

shirking condition faced by employers the optimal contracts could potentially imply work hours which are lower than the first-best optimum. However, the opposite could also very well be the case. The analysis by Rebitzer and Taylor (1995) thus has the following implication:

Proposition 7: *The pooling equilibrium may involve a length of the work day which exceeds the first-best optimal choice of hours. In particular, it may very well be the case that the short-hour type workers must work more than what is socially optimal*

Even though all firms would benefit from a simultaneous switch to the separating contract such an allocation is unlikely to be sustained in equilibrium. In fact, Rebitzer and Taylor show that for the relevant range of θ there exists no pure-strategy equilibrium in which all firms offer the short-hour jobs associated with the separating contracts. Their analysis thus demonstrates that the market forces may cause the provision of short-hour jobs accommodating individuals with preferences for more leisure may be inefficiently low. They do argue, however, that such an allocation could be sustained through the passing of legislation forcing firms to also offer contracts with shorter hours.

3.3 Unobservability of agents' talent and productivity

The model presented in Landers et al. (1996) and discussed above is one in which agents have *career concerns*. Even attorneys with preferences for shorter work hours would be interested in a career as partner in a law firm if this automatically would imply an increase in hours compensation (due to both higher productivity in such a position as well as the equal sharing of profits amongst partners). Since the current partners may not always be able to correctly infer the value of the contribution of the prospective partners to the firm, even associates with preferences for short hours may have an incentive to work long hours early in their career in order to be promoted as partner. In order to avoid such a situation, the firm will require its associates to work beyond their Pareto efficient level, to the point where attorneys with preferences for shorter hours refrain from applying to these positions. Thus the career concerns together with the incomplete information are the source of an economic inefficiency.

The idea of an inefficiency associated with career concerns was first put forward by Holmström (1982). Again, the inefficiency is a result of incomplete information including

unobservability of agents' characteristics. However, contrary to the model in Landers et al. (1996) in which agents differ in their preferences for leisure, Holmström considers a model in which agents' productivity is not known to employers. In addition, employers are not able to observe agents' input of effort in production, and thus the model also includes moral hazard. The following exposition is based on Bolton and Dewatripont (2005):

We consider a simple two-period model. Output in each period is given by

$$q_t = \theta + e_t + \omega_t, \quad t = 1, 2. \quad (31)$$

Here θ represents the agent's talent, e_t is the agent's effort level and ω_t a random shock to production in period t . Neither of these variables can be directly observed by the employer. However, the (prior) probability distribution of both θ and ω_t is known to all parties. We assume that the agent is risk-neutral, having period-utility represented by

$$u(y_t, e_t) = y_t - c(e_t), \quad (32)$$

where y_t denotes the agent's compensation in period t and the function c represents the agent's disutility of effort. As in earlier models, we assume that this function satisfies $c(0) = 0$, and $c'(\cdot), c''(\cdot) > 0$. Let β denote the agent's discount factor.

The realisation of output is observed by everyone in the market. However, it is assumed that contracts cannot be written contingent on the realisation of the output in the same period. Also, no multi-period contracts are allowed. Hence, employers offer fixed wages in every period. In the first period, as agents' productive abilities are unknown, all agents are offered the same compensation y_1 . However, in the second period, employers compete for the employees with the highest talent. Only the realised output in period one can be used to make inference about the agent's abilities, and thus the second period wage will be a function of output in period one, i.e. $y_2(q_1)$. Competition among employers in period two makes the second period compensation equal the expected output in that period. It is immediately obvious that the agent will not have any incentive to exert effort in the second period. Hence the expected output in the second period will equal the market's belief about the agent's ability, denoted by θ^* . Let e_1^* be the market's belief about the effort exerted by the agent in the first period. Then the expected output in the second period, which equals the second-period compensation, is given by

$$y_2(q_1) = E(\theta | q_1) = \theta^* = q_1 - e_1^* = \theta + e_1 - e_1^* \quad (33)$$

Given a fixed wage y_1 in period one, and the fact that the agent will choose to exert zero effort in the second period, the agent's lifetime utility can be written as

$$y_1 - c(e_1) + \beta[\theta + e_1 - e_1^*] \quad (34)$$

As Bolton and Dewatripont (2005) point out, we have the following result:

Proposition 8: *The agent's first-period input of effort exceeds the social optimum if and only if $\beta > 1$.*

To see that this must hold note that the agent maximises his lifetime utility with respect to his first-period input of effort. The first-order condition solving this problem is given by

$$c'(e_1) = \beta \quad (35)$$

On the other hand, efficiency would require an equal allocation of effort in both periods, with the first-best period input of effort, e_t^{FB} , satisfying the equation

$$c'(e_t^{FB}) = 1 \quad (36)$$

for $t = 1, 2$. Obviously, the market outcome yields a second-period effort level which is necessarily too low. For the first-period effort the conclusion depends on the size of the agent's discount factor and since the marginal disutility of effort is increasing we have

$$e_1 \geq e_t^{FB} \Leftrightarrow c'(e_1) \geq c'(e_t^{FB}) \Leftrightarrow \beta \geq 1, \quad (37)$$

and the above Proposition follows. If $\beta > 1$ the agent's private marginal benefit of exerting effort in the first period, thus improving the prospects of higher pay in the future, exceeds the social benefit and the agent thus finds it optimal to increase his labour input beyond the socially optimal level.

It may seem strange to consider the latter case where the discount factor exceeds unity. However, this may be a natural assumption if the second period is much longer than the first. An example of this would be the system of tenure associated with positions in academia: Students and post-doctoral researchers may have an incentive to exert high levels of effort early on in their career in order to get tenure at a later point. Knowing this to be the case, even the more able students, who objectively are academically more fit for the position will be forced to work excessively much in order to ensure these positions. At the end though, one might hope that only the most able end up in these positions.

The present model exhibits some of the same features: Even the agents who observe that their talent is above normal will have to exert high effort in the first period in order to get a "decent" pay in the second period. The reason is that the market anticipates that agents will exert the high effort levels in the first period in order to boost their second-period compensation. In equilibrium, thus, $e_1^* = e_1$ and the agent's second-period expected compensation only equals the true productivity θ .

Holmström's original model involved an infinite number of periods. The conclusions are somewhat less extreme than in the above version. In particular, the agent will continue to exert positive effort at a stationary level as long as the market is unable to correctly infer the agent's true talent. However, this stationary level never exceeds the socially efficient level. He finds that agents are thus usually prone to exert higher effort early on in their career. He writes:

Normally, we expect that the precision of information about ability increases as time goes on. [I]n that case young people will overinvest in labour supply because the returns from building a reputation are highest when the market information is most diffuse.

Career concerns, its effect on optimal incentive schemes, as well as its empirical relevance, are further discussed by Gibbons and Murphy (1992), Dewatripont et al. (1999 a,b) and Kaarbøe and Olsen (2006), among others.

The focal point of this section has been the analysis of the determination of working hours and effort in a context of imperfect information. We have discussed several contributions to the literature establishing that imperfect information may lead to a market outcome in which agents' work hours or effort level is inefficiently high. We started by noting that in a setting of pure moral hazard, that is, when only the agent's input of effort is unobservable to employers, there exist parameter values such that the equilibrium allocation involves over-provision of effort, provided that the agent is risk-averse. We next considered several models where one or more characteristics of individuals are not directly observable. We have showed that with such population heterogeneity together with incomplete information the prospects of adverse selection can easily lead to Pareto inefficient outcomes where individuals work too much. The information asymmetry leads to a situation in which the private return from increasing work effort exceeds the social

benefit. To avoid a situation where individuals with undesirable characteristics fill a given position, the solution may be to increase the required length of the work day beyond the Pareto efficient level so as to make the position undesirable for individuals with such characteristics. The result may be that the provision of short-hour jobs in the market is sub-optimally low. When effort cannot be contracted upon, individuals may also find it optimal to exert high levels of effort, in particular if this entails higher compensation in the future. Thus an appropriate mix of adverse selection, moral hazard and dynamics can very well lead to Pareto inefficient market outcomes due to individual overemployment.

4 Expanding the literature on individual overemployment

In the previous two Sections we have presented and discussed some of the most central contributions to the literature in which individual overemployment occurs as a market inefficiency. We have seen that two of the classical sources of market inefficiency, namely the presence of externalities and the lack of perfect information, easily can be used to construct equilibria characterised by individual overemployment. In this Section we present two additional models in which individual overemployment occurs as a result of imperfect information. Contrary to some of the previous models, our models will simultaneously exhibit both overemployment *and* underemployment.

We first present a simple model of adverse selection which is motivated by a concern regarding the realism of a central assumption made in several of the models of the previous section. Upon relaxing this assumption we are still able to construct an example where overemployment presents itself. Thereafter we develop a variation of a shirking-model inspired by one of the models presented in the previous section. Over model, however, is more in line with the idea presented in Shapiro and Stiglitz (1984), where aggregate unemployment is necessary to prevent workers from shirking. We extend the model by including individual work hours as an additional variable. We find that under certain conditions the market equilibrium length of the work-day exceeds the Pareto optimal level.

4.1 A simple model of adverse selection in teams

In Section 3 we noted the interjection that the model presented in Akerlof (1976) was based on a set of peculiar assumptions which made several commentators question the empirical relevance of the paper's subsequent discussion and description of the model's equilibrium. For instance, the assumption that agents with higher talent/productivity also have lower disutility of effort may seem constructed with the sole purpose to obtain an equilibrium characterised by individual overemployment, although one perhaps could find cases where such an assumption may be reasonably realistic. Partly for fear that the issue of overemployment should not be taken seriously, some authors thus sought to construct other models in which individuals would end up working inefficiently much under less stringent assumptions. In this context many of the career concern models represent an interesting direction of research, as these models often produce equilibria with both under- and overemployment. In particular, a common feature of these models is that agents are inclined to work too much early in their career but exert too little effort later on.

The career concern models also frequently make use of additive production functions, such as the one used in our discussion of Holmström's original model in Section 3.3. If we disregard any stochastic elements, the level of output, q , could then be expressed by the simple relation

$$q = \theta + h, \tag{38}$$

where θ denotes the agent's talent or productive abilities, and h his disutility-generating input into production, such as effort or work hours. Such an additive form was also used in the model by Akerlof (1976). Although this functional form features the plausible notion that more talented workers produce more than others for any given level of the input h , the fact that output will be strictly positive even when no input of effort or work hours are used may seem somewhat less realistic. On the other hand, this is one of the driving forces of the mechanism in the career concerns models leading to the result that agents might exert too high levels of effort early in their life, as we saw in Section 3.3. It could be argued that such a production function is still quite realistic if we implicitly assume that the firm has access to some sort of imperfect monitoring technology. In this case, θ could be interpreted as the output resulting from the agent's effort that the employer is able to induce due to monitoring. As such, h could be thought of as the agent's contribution to

production which cannot be monitored by the employer. Still, without explicitly modelling such factors, the above specification of the production technology may seem quite absurd. Therefore, it would be of great interest to examine whether equilibria characterised by individual overemployment can still be described when we allow more commonly used, and perhaps more plausible production functions.

We consider here a model based on the idea in Akerlof (1976) that agents work in teams. Specifically, we assume that production takes place with employees working in pairs. Only when both workers are present will output be produced. As a consequence, both members of the team will work the same amount of hours, h . Also similar to Akerlof (1976), agents have different productive ability, with θ_i denoting the ability of worker i . For simplicity we assume that this talent can take one of two values, i.e. $\theta_i \in \{a, b\}$, with $a > b > 0$. Contrary to Akerlof, however, we assume that the agents' talent affects the teams *marginal* productivity. For ease of calculation, we consider the case where marginal productivity is constant. Specifically, we will assume that the team's output is given by

$$q = (\theta_1 + \theta_2)h \quad (39)$$

There is a large number $N \in 2\mathbb{N}$ of agents in the population, with a fraction $p \in (0, 1)$ of agents having marginal productivity $\theta_i = a$ and the rest having marginal productivity $\theta_i = b$. Agents are assumed to be risk-neutral with increasing marginal disutility of labour. For simplicity, we will assume that agents' preferences are given by

$$U(y, h) = y - \frac{1}{2}h^2, \quad (40)$$

where y denotes the compensation received. Note that all agents have equal preferences, contrary to what was assumed in Akerlof (1976). However, similarly to Akerlof we will assume that only agents know their respective marginal productivity, that is, θ_i can neither be observed by firms nor other agents before production begins. Furthermore, we assume that the market is characterised by perfect competition, meaning that firms will earn zero expected profits in equilibrium. To make matters easy we will also assume that the compensation paid to workers must be contracted upon and paid before production takes place. Clearly, since all agents have the same preferences, an individual firm cannot suggest a wage-hour package which only attracts workers of a specific type. Also, since any firm cannot know what type of workers are being employed by other firms, it follows that, to the extent that the firm is able to attract workers at all, with probability p any

of the two workers will have marginal productivity a , and with probability $1 - p$ either of the two will have productivity b . Hence, expected profits will be given by

$$E(q) = 2[pa + (1 - p)b - y] \quad (41)$$

As noted above, competition among firms cause equilibrium profits to be zero. Meanwhile, as firms seek to attract workers, they construct the wage-hour package that maximises agents' utility, subject to having non-negative profits. Hence the firms' equilibrium optimal contract will specify a length of the work day, h^* , given by

$$h^* := \operatorname{argmax}_h \left\{ [pa + (1 - p)b]h - \frac{1}{2}h^2 \right\} = pa + (1 - p)b, \quad (42)$$

and consequently an equal pay to both workers amounting to $y^* = [pa + (1 - p)b]^2$. Thus, workers' equilibrium utility, U^{eq} , will reach a level of:

$$U^{eq} = \frac{1}{2}[pa + (1 - p)b]^2 \quad (43)$$

The main point of this model is that the market outcome is inefficient and that welfare could increase if some agents were to work fewer hours, and others were to work longer hours, very much like we saw in the career concerns model examined in Section 3.3. We have the following result:

Proposition 9: *The market outcome entails two inefficiencies: The fraction p of a -type individuals in the population works too little, whereas the fraction $1 - p$ of b -type individuals works too much. Furthermore, given the information asymmetry, there exists a policy under which an equilibrium with the efficient allocation of work hours for both types can be sustained.*

Proof: It is easy to establish that the optimal production scheme would pair up workers of the same ability: Consider two pairs of agents, one of each type. Since agents are risk-neutral and we allow for side payments the optimal allocation would simply maximise the net social benefits of production. Obviously, if agents of the same type are put together in teams, the optimal work hours would be given by θ , and the social benefit associated with production would equal $2[\theta^2 - 1/2\theta^2] = \theta^2$ for this pair. If we rather put agents of different type together, the optimal amount of work hours would be $(a + b)/2$, and the total social benefit of production would be given by $4[(\frac{a+b}{2})^2 - \frac{1}{2}(\frac{a+b}{2})^2] = \frac{1}{2}(a + b)^2$. Now, the former allocation clearly yields a higher social benefit, since

$$a^2 + b^2 - \frac{1}{2}(a + b)^2 = \frac{1}{2}[a^2 + b^2 - 2ab] = \frac{1}{2}(a - b)^2 > 0 \quad (44)$$

Hence, the efficient allocation has agents of equal type work together in teams. Now, even in the presence of our information asymmetry, there exists an equilibrium in which the efficient allocation can be sustained under an appropriate choice of policy. The crucial requirement is that the agent's utility from taking either job cannot depend on his type. Let t denote an income tax for individuals working in a -teams, which will equal the wage subsidy received by members of a b -team. The tax/subsidy equalises the utility from working either job, meaning that

$$\begin{aligned}\frac{1}{2}a^2 - t &= \frac{1}{2}b^2 + t \\ \Leftrightarrow t &= \frac{1}{4}(a^2 - b^2)\end{aligned}\tag{45}$$

Of course, the tax/subsidy scheme must be financed, which can be done by introducing a lump-sum transfer s for all agents satisfying

$$s = pt - (1 - p)t = (2p - 1)t = \frac{1}{4}(2p - 1)(a^2 - b^2),\tag{46}$$

which can be both positive and negative, depending on the proportion p of type a individuals in the economy. Consequently, the utility, U^{int} , from working either job following the market intervention, using that $p \in (0, 1)$ and $a > b$, will satisfy

$$\begin{aligned}U^{int} &= \frac{1}{2}a^2 - t + s = \frac{1}{4}(a^2 + b^2) + \frac{1}{4}(2p - 1)(a^2 - b^2) \\ &= \frac{p}{2}(a^2 - b^2) + \frac{1}{2}b^2 = \frac{p}{2}(a - b)^2 + pb(a - b) + \frac{1}{2}b^2 \\ &> \frac{p^2}{2}(a - b)^2 + pb(a - b) + \frac{1}{2}b^2 = \frac{1}{2}[p(a - b) + b]^2 \\ &= \frac{1}{2}[pa + (1 - p)b]^2 = U^{eq},\end{aligned}\tag{47}$$

that is, all agents are better off than under the free-market equilibrium. In this Pareto optimum, a fraction p of all agents work more than in the former equilibrium, thus eliminating part of the inefficiency of the latter. However, a fraction $1 - p$ now also works less than in the free-market outcome, which straightens out the second inefficiency of the market allocation. \square

The above model is a simple example successfully establishing the existence of a pure adverse selection model where the market outcome is inefficient, partly because a fraction of the population works more than the Pareto efficient level. Furthermore, this result was obtained using a production function which does not exhibit the rather implausible feature

that production can be positive even when labour input is zero. In the next subsection we develop a model with moral hazard in which individual overemployment is also a possible outcome of the market equilibrium.

4.2 An extension of the Shapiro-Stiglitz model

4.2.1 Unemployment and the inadequacy of the Walrasian model

A major drawback of the perfect competition model and its many extensions (such as the one analysed in Section 2) is that it does not account for involuntary unemployment. In the representative-agent paradigm, in particular, the entire population is represented by a single individual, who optimally chooses his personal labour supply as a response to the prices in the market, which he takes as given. As such, unemployment is in some sense absent from the standard Walrasian theories of microeconomics. Meanwhile, unemployment is a prominent feature in the Keynesian paradigm of Economics, in which involuntary unemployment occurs as a consequence of insufficient demand for labour and the failure of prices to adjust so as to clear the market. In its basic form, however, this approach does not explain the origins of the unemployment, but rather simply ascribes it to an assumed (temporary) nominal price rigidity, which is seen as one of the fundamental characteristics of the economic environment. This view therefore strongly conflicts with the idea of a Walrasian equilibrium. Thus Weitzman (1982) writes:

Unemployment equilibrium as a story about effective demand failure co-exists uneasily with classical general equilibrium theory. Most economists deal with both concepts, but in a kind of schizophrenic manner. Unemployment and classical versions of equilibrium theory are not just different approaches; in a fundamental sense they seem to represent almost incompatible paradigms.

The shortcomings of both these approaches prompted researchers to create a new microeconomic foundation suited to produce equilibria characterised by involuntary unemployment, that is, equilibria where some individuals end up without a job, but are still willing to work at the prevailing market wage rate.

The subsequent array of research took many different directions and brought forth many important insights into the effects of unemployment on the fluctuations of production in the economy. Many of these models also provided predictions which were more successful

in fitting the available data on the economy's behaviour in the course of the business cycle. One such contribution were the works by Hansen (1985) and Rogerson (1988) on *indivisible labour*, which

is modeled by assuming that individuals can either work some given positive number of hours or not at all - they are unable to work an intermediate number of hours. This assumption is motivated by the observation that people either work full time or not at all. (Hansen (1985))

Another strain of research which successfully incorporated the idea of unemployment was the soon to be extensive field on *search and matching* models, developed by Mortensen (1970), Lucas and Prescott (1974), Diamond and Maskin (1979), Pissarides (1979), Burdett and Mortensen (1980) and Diamond (1982), to mention some.

Much of this early work ignored the determination of the length of the work day, which was exogenously given. These models were therefore neither fit to confirm nor rebut the normative conclusions of the classical Walrasian model pertaining to the market efficiency regarding work hours. Being able to explain the existence of unemployment, many of these models also provided a framework within which unemployment not necessarily be viewed as a bad thing, although many of these models also draw the conclusion that equilibrium unemployment may be inefficiently high.

Meanwhile the advances in the field of non-cooperative game theory enabled the development of the principal-agent framework and contract theory, which allowed for the analysis of the determinants of individual work hours. A substantial part of this literature drew attention to the role of information asymmetries in the contractual framework between the principal and the agent, and its effect on the economic outcome. A conclusion not uncommon in much of this work would be that the delegation of tasks in a context where the agent's effort level cannot be perfectly observed, often leads to a situation where the latter exerts effort at a level which is inefficiently low. As we argued in Section 3, the emphasis on this conclusion may be a result of researchers' bias towards simplifying the analysis and restricting attention to the case where both parties are risk-neutral.

Alongside the development of the principal-agent framework yet another class of models rationalising the observed wage rigidity in the data was developed. These were the *efficiency-wage* models, which recognised the potential positive effect of higher wages on the workers' output level. In its crudest form, this view could be motivated by the idea

that workers who were able to enjoy a higher consumption also easier developed the necessary strength to contribute to the production process. If wage-setting firms take this effect into account, they may be inclined to set the wage above the market clearing wage of the Walrasian model. Under the assumption that all workers must be paid the same wage, such models then easily result in equilibria where involuntary unemployment occurs: At the market wage rate firms employ an optimal number of workers and are therefore not interested in hiring anyone else. And even though there are individuals who would be willing to work at a lower wage than the market rate, firms have reservations about decreasing the wage as this would also decrease the productivity of their entire work stock. Note that this argument hinges on the assumption that wages cannot be differentiated across workers. See Yellen (1984) for a survey of the efficiency-wage approach to unemployment.

Although the story about the worker, who as a consequence of better nutrition manifests increased productivity, may sound plausible in certain economic contexts, it fails to be a credible explanation of the existence of unemployment. In particular, it does not explain why firms would be reluctant to hire extra workers at a lower wage than the other employees. To deal with this issue, the efficiency-wage approach drew on the insights from the work on incentive problems and imperfect observability of employees actions. This resulted in the shirking-models developed by Bowles (1981) and Stoft (1982), with Shapiro and Stiglitz (1984) being the most important contribution. In these models unemployment would be considered necessary to induce workers to exert the desirable level of effort: If exerting effort is costly to the employees, and there is a positive probability that the individual effort level not be observed by the employer, then workers might have an incentive not to exert effort, that is, to shirk. In order to induce workers to exert effort the fear of losing one's job when caught shirking must be so high that exerting effort is the only viable option. This idea better motivates the efficiency-wage approach: In this framework employers might find it optimal to pay high wages so as to increase the loss incurred by employees if losing their job. This, however, also requires the unemployment rate to be sufficiently high so that current employees know that finding a new job after being laid off will be difficult. Moreover, employers do not have an incentive to hire additional workers at a lower wage rate, seen as these might not be encouraged to exert the required effort level and thus will represent a pure cost to the firm. It goes without saying

that unemployment in this case represents an economic inefficiency.

4.2.2 Individual work hours in the context of aggregate unemployment

The original model by Shapiro and Stiglitz (1984) took the workers' effort level to be a binary variable. They noted in their footnote 4 that "[i]ncluding effort as a continuous variable would not change the qualitative results". With regards to generating equilibrium unemployment this may be true. However, as we shall see, the unobservability of effort may also be the source of a second form of inefficiency, pertaining to the determination of individual work hours in the economy. As a consequence, the one-sided emphasis on underemployment in their model may be a result of restricting attention to the case where an individual's contribution to the production process is solely determined by a binary variable.

In the succeeding discussion we will follow the idea in Rebitzer and Taylor (1995), who present a model analysing the determination of working hours stipulated in the contracts offered by firms. However, they choose not to explicitly include unemployment as a variable of their model. In addition, it is my contention that their analysis (which restricts attention to cost minimisation) blurs the potential implications of firms' employment level on the determination of individual work hours. To keep the discussion closer to the idea in Shapiro and Stiglitz (1984) we will consider firms' profit maximisation problem, and also include movements in and out of employment, with unemployment serving as a "discipline device".

Consider an economy with a continuum of length N of identical firms, all producing a homogeneous output employing a production technology represented by the function F . The level of output in any given period depends on the effective work hours by the firm's employees. The work effort of an individual worker cannot be perfectly observed. As in Shapiro and Stiglitz (1984) we take the effort level to be binary, meaning that workers may choose to either exert a positive level of effort, or no effort at all. We let the number of employees who exert effort in a given period be L_t and the number of workers who choose to shirk be S_t . All employees work a given number of hours h_t , which can be monitored. Shirking workers do not contribute to the production level. Thus the output of the firm is given by $F(L_t h_t)$. As is common in the literature, we assume that $F'(\cdot) > 0$ and $F''(\cdot) < 0$.

We assume that the economy is composed of a continuum $N\bar{L}$ of identical and infinitely-lived individuals. In any given period t , an individual may find itself in one of three states; being employed and exerting effort, being employed while shirking, or being unemployed. Let the period utility level of being in either state be denoted by U_E , U_S and U_U respectively. Regardless of the wage-work hours package offered by firms, the problem is that we always have $U_S > U_E$, meaning that workers prefer not to exert effort. To simplify the below analysis we consider the case where individuals' utility is quasi-linear. That is, we assume that preferences are represented by

$$U_i(y, h) = y - c_i(h), \quad (48)$$

with $c_i(0) = 0$ for $i = E, S$. Now a worker's wish to shirk stems from the fact that the disutility from working a given number of hours is higher for employees exerting effort, that is $c_E(h) > c_S(h) \forall h > 0$. We further assume that the disutility of further hours is a convex twice differentiable function satisfying $c''_i(\cdot) > 0$ and $c'_E(h) > c'_S(h) > 0$.

To include movements in the labour force in and out of employment also in the steady-state, we assume that job breakups occur with a constant probability $q \in (0, 1)$, that is, a position is terminated and the former employee becomes unemployed. Furthermore, with probability D shirking workers are detected and let go. As was assumed in Rebitzer and Taylor (1995), this probability is a function of the number of hours worked by the employee. We follow this idea and assume that the function is differentiable, satisfying $D'(h_t) > 0$. The idea is that employers may use the length of the work-day as an improvement on his possibility of monitoring the workers' effort: If there is a positive probability that shirking workers will be undetected at any given point in time, it may be reasonable that the probability they be detected during the entire work-day is higher, the longer is the work-day. Finally, let $r > 0$ denote the individuals' discount rate.

The dynamics are as follows: Upon the completion of production in any period t , all employees are paid a compensation y_t for their work in this period. (This assumption implies that also workers caught shirking will be paid, which is reasonable if the individual work effort cannot be observed by a third party and/or if it is costly to the firm to involve the judicial system in disputes over effort levels.) The workers all consume their compensation in period t , thus there is no saving. Output is produced, with all employees working h_t hours. At the end of the work-day any workers caught shirking during the day will be fired and any other position is terminated with probability q . At the beginning of

next period, firms search to replace the terminated employees and any of the unemployed gets a new job with probability p_{t+1} . The firm then proposes a contract to all employees, which they either accept or reject. Upon rejection, an individual will have to wait until the next period before finding a new job. Similarly, the firm will have to wait until next period before it can replace the worker(s) who rejected the proposed contract. For simplicity, we will also assume that it is impossible to write contracts involving a compensation scheme over several periods. With these assumptions about the framework within which the contracting takes place, it follows that in a steady-state equilibrium all firms will offer the same contract to every employee.

In the following we are looking for the steady-state behaviour of the economy and therefore drop the time subscript of all variables. Included in the assumption of a steady-state is also the idea that a worker who finds it optimal to shirk in any given period will continue to do so in any subsequent period. This follows from the fact that individuals are infinitely-lived and the fact that in a steady-state the future will look the same regardless of the time subscript and the history of events. Similarly, if it is ever optimal to exert effort in the steady state, it will continue to be so. We therefore distinguish between two states of unemployment; one where a worker exerts effort once he gets a job, UE , and one where the workers shirks, US . In either case, we assume that the unemployed receive zero period utility.

Now let V_i denote the expected discounted present lifetime value of being in state $i = E, S, UE, US$. Consider a worker being in state E , that is, a worker who exerts effort. In the present period he receives a utility level of U_E . With probability $1 - q$ his job is not terminated and he continues in the job, thus anew receiving the lifetime value V_E , whose value he discounts to the present period. With probability q the worker's job is terminated. However, with probability p he already finds a new job at the beginning of the next period, in which case he finds himself in the same situation as in the present. Lastly, if his job is terminated, there is a probability $1 - p$ that he remains unemployed in the next period, thus earning the value of being an unemployed, but potentially effort-exerting worker. It follows that V_E can be written as

$$V_E = U_E + \frac{1}{1+r}[(1-q+qp)V_E + q(1-p)V_{UE}] \quad (49)$$

From the fact that employees who have once found it optimal to exert effort once we have reached a steady state will continue to do so if they get a new job, we can express V_{UE} in

the following way: In the current period as unemployed the worker receives nothing, but with probability p he will get a job in the next period and thus be in state E . Otherwise, with probability $1 - p$ he will continue to be unemployed in the next period. We therefore have

$$V_{UE} = \frac{1}{1+r} [pV_E + (1-p)V_{UE}] \quad (50)$$

Solving for V_{UE} , inserting into the above equation and rearranging, we find that the expected discounted present lifetime value of being an employed worker exerting effort can be written as

$$V_E = \frac{(p+r)(1+r)}{r(r+q(1-p)+p)} U_E \quad (51)$$

Similar calculations can be done to find an expression for V_S : In the current period a shirking worker receives a utility level of U_S . With probability $(1-q)(1-D)$ he is neither caught shirking, nor is his position being terminated, and he therefore continues to shirk in the next period. With probability $(1-q)D$ the position is not being terminated, but the worker is caught shirking and therefore let go and with probability q the position is simply terminated. In both these cases the worker will, with probability p find a new job in the next period and continue to shirk, and with probability $1-p$ he will be unemployed in the next period. Hence we can write

$$V_S = U_S + \frac{1}{1+r} (1-q)(1-D)V_S + [(1-q)D + q](pV_S + (1-p)V_{US}) \quad (52)$$

Analogous to the above result, we find that

$$V_{US} = \frac{1}{1+r} [pV_S + (1-p)V_{US}] \quad (53)$$

which, if inserted into the previous equation yields, upon rearranging, the following expression for V_S :

$$V_S = \frac{(1+r)(p+r)}{r(p+r+(1-p)[D+q(1-D)])} U_S \quad (54)$$

Now, the fundamental lesson to be learned from the efficiency-wage/shirking literature is that rational employees will only choose to exert effort if the expected utility from doing so exceeds that of shirking, i.e. the no-shirking condition states that $V_E \geq V_S$, which can be rewritten as

$$U_E \geq \frac{p+r+q(1-p)}{p+r+(1-p)[D+q(1-D)]} U_S \quad (55)$$

From this equation it is immediately clear that as long as the probability of finding a new job in the next period, upon being fired due to shirking, is sufficiently close to one, then

no worker will have an incentive to exert effort at all. At the same time, firms will be aware of this fact, and will not be willing to hire anyone as long as the probability is at this level. Now, the probability of finding a new job is, in the present model, inversely related to the unemployment rate. It follows that only when the unemployment rate is sufficiently high will workers have an incentive to exert effort and firms be willing to employ these workers. This is the qualitative conclusion of Shapiro and Stiglitz (1984): Employees' incentive to shirk and the employers' inability to observe the workers' effort level requires the existence of positive equilibrium unemployment in order for firms to be willing to employ anyone at all. Unemployment works as a "discipline device". However, it should be clear that such an outcome is necessarily Pareto inefficient. This therefore warrants the use of the term *underemployment*.

We now turn to the determination of individual work hours. We first note that the no-shirking condition can be rewritten as

$$U_E \geq \frac{p + r + q(1 - p)}{(1 - p)(1 - q)D}(U_S - U_E) > 0 \quad (56)$$

Inserting the specific functional forms of U_i from (48) into (56) we find that the no-shirking condition can be written as

$$y - c_E(h) \geq \frac{p + r + q(1 - p)}{(1 - p)(1 - q)D(h)}(c_E(h) - c_S(h)) \quad (57)$$

In a steady-state where all employees exert effort the period profits earned by firms will be given by

$$\pi = F(hL) - yL \quad (58)$$

It is immediately obvious that if the no-shirking condition were not binding the firm could easily increase its profits by simply decreasing y . Hence in optimum, (57) must be binding. This stems from the fact that the firm does not face any difficulties in attracting labour. Given that all other firms propose a contract where (57) binds, the only sub-game perfect Nash equilibrium would be for any individual firm to propose the same contract. As in Shapiro and Stiglitz (1984) we will assume that labour be efficiently allocated among firms, meaning that all firms employ the equal number of workers (see their footnote 9). This is consistent with equilibrium.

Define $C(h) := c_E(h) - c_S(h)$. Now the firm's maximisation problem consists of finding the optimal choice of L and h . Let these optimal values be denoted by L^* and h^* respectively. Assuming that an inner solution to the firms' profit maximisation problem

exists, these variables will satisfy the following equations (derived from the first order conditions):

$$h^* F'(h^* L^*) = y = c_E(h^*) + \frac{p + r + q(1 - p)}{(1 - p)(1 - q)} \frac{D(h^*)}{C(h^*)} \quad (59)$$

$$F'(h^* L^*) = c'_E(h^*) + \frac{p + r + q(1 - p)}{(1 - p)(1 - q)} \frac{C(h^*)}{D(h^*)} \left[\frac{C'(h^*)}{C(h^*)} - \frac{D'(h^*)}{D(h^*)} \right] \quad (60)$$

Note that at the macroeconomic level, the aggregate demand for labour pushes the probability p upwards, thus tightening the no-shirking condition, which ensures that no firm will have an incentive to increase their employment in the steady-state equilibrium. Now the latter equation corresponds to equation (6) in Rebitzer and Taylor (1995). They discuss the efficiency properties of the contracting outcome as a function of the sign of the second term on the right-hand side of (60). The basic conclusion for the model that they consider is that the contracted work hours are inefficiently low or high when the second term is positive or negative respectively. As they note either case is in principle possible, as is the situation where the second term is zero, in which case the contracted length of the work-day would be Pareto efficient. It should also be noted that an inner solution to the problem requires that the second term on the right-hand side of (60) not be too negative, otherwise the first-order condition could imply $F'(h^* L^*) < 0$ which is impossible. In the below discussion we assume that an inner solution to the problem always exists.

As we have already insinuated, the above conclusions are in general no longer valid under the present assumption that individual productivity depends on the firm's overall employment level. With decreasing marginal productivity of man-hours, the efficiency merits of the firms' determination of individual work hours depend not only on the sign of the second term on the right-hand side in (60). We have the following result:

Proposition 10: *Let h_{FB} denote the first-best optimal amount of work hours. The contracted length of the work-day, h^* , exceeds h_{FB} if and only if*

$$\frac{p + r + q(1 - p)}{(1 - p)(1 - q)} \frac{C(h^*)}{D(h^*)} \left[\frac{C'(h^*)}{C(h^*)} - \frac{D'(h^*)}{D(h^*)} \right] < F'(h^* L^*) - F'(h_{FB} \bar{L}) \quad (61)$$

Proof: It is easy to verify that a first-best Pareto efficient allocation would have to satisfy full employment of labour equally distributed among firms with optimal individual work-hours satisfying the condition

$$F'(h_{FB} \bar{L}) = c'_E(h_{FB}) \quad (62)$$

Using this together with (60) we now have that

$$\begin{aligned}
& \frac{p+r+q(1-p)}{(1-p)(1-q)} \frac{C(h^*)}{D(h^*)} \left[\frac{C'(h^*)}{C(h^*)} - \frac{D'(h^*)}{D(h^*)} \right] < F'(h^*L^*) - F'(h_{FB}\bar{L}) \\
& \Leftrightarrow F'(h^*L^*) = c'_E(h^*) + \frac{p+r+q(1-p)}{(1-p)(1-q)} \frac{C(h^*)}{D(h^*)} \left[\frac{C'(h^*)}{C(h^*)} - \frac{D'(h^*)}{D(h^*)} \right] \\
& \quad < c'_E(h^*) + F'(h^*L^*) - F'(h_{FB}\bar{L}) \\
& \Leftrightarrow c'_E(h_{FB}) = F'(h_{FB}\bar{L}) < c'_E(h^*) \\
& \Leftrightarrow h_{FB} < h^*,
\end{aligned} \tag{63}$$

where the last inequality uses that c'_E is increasing in its argument. \square

The above result may seem somewhat tautological at first, and its main intuition may perhaps be better expressed by the following:

Corollary 2: *Assume that*

$$\frac{p+r+q(1-p)}{(1-p)(1-q)} \frac{C(h^*)}{D(h^*)} \left[\frac{C'(h^*)}{C(h^*)} - \frac{D'(h^*)}{D(h^*)} \right] \leq 0 \tag{64}$$

Then the market equilibrium length of the work-day exceeds the first-best optimal h_{FB} .

Proof: From the fact that the equilibrium requires strictly positive unemployment (and under the assumption of efficient allocation of labour among firms) it follows that the individual firm's privately optimal level of labour is less than the efficient one, i.e. $L^* < \bar{L}$. This means that for any level of h would $F'(hL^*) > F'(h\bar{L})$. As a consequence, the right-hand side of (61) can only be negative if $h^* > h_{FB}$ and thus the result follows from the above Proposition. \square

From the idea in the above proof it should be clear that the contracted work hours may exceed the first-best solution even when the second term on the right-hand side of (60) is positive. Thus the model may simultaneously involve underemployment, in the sense that firms choose to hire too few workers, as well as overemployment in the sense that the employed workers exert themselves too much. The fact that the steady state market allocation of individual work hours exceeds the first-best optimum in the case where the second term on the right-hand side of (60) is zero is illustrated in Figure 2.

It could be argued that the market solution still represents a second-best solution. That is, given the equilibrium unemployment rate, in the case where $\frac{C'(h^*)}{C(h^*)} - \frac{D'(h^*)}{D(h^*)} = 0$ the employees would still be working the Pareto efficient level. This is of course true, however

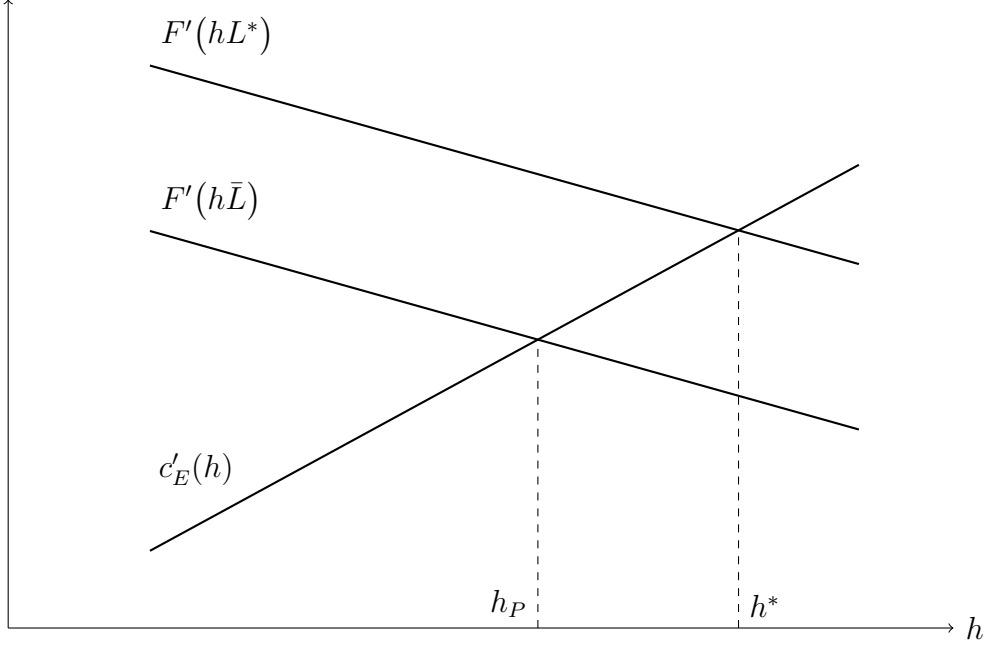


Figure 2: The contracted optimal hours of work exceeds the Pareto efficient level under the assumption that $\frac{C'(h)}{C(h)} - \frac{D'(h)}{D(h)} = 0$

this should still only be considered a third-best analysis: The relevant restriction to consider is the no-shirking condition, and under this constraint the market outcome is not in general efficient. This is demonstrated in Shapiro and Stiglitz (1984) where the level of employment, L , and worker compensation y , are the only variables under consideration. They find that, taking into account the information costs, the aggregate unemployment level is still too high. The second-best inefficiency of the market outcome translates to our extended model. However, under certain conditions, we also find that individual work hours may be too long:

Proposition 11: *Assume that $\frac{C'(h^*)}{C(h^*)} - \frac{D'(h^*)}{D(h^*)} = 0 \forall h$, meaning that the right-hand side of the no-shirking condition in (57) does not depend on h .⁴ Then the steady state market equilibrium determination of employment is sub-optimally low, and the individual work hours are sub-optimally long, even when the information costs are taken into account.*

We consider the case where every member of the population $N\bar{L}$ receives an equal share of the aggregate profits. Under the assumption that workers are risk-neutral, this would not change the no-shirking condition in (57). The social planner's problem is to maximise the

⁴This would be the case if both the functions $C(h)$ and $D(h)$ are linear in h .

sum of the $N\bar{L}$ individuals' utility. However, as Shapiro and Stiglitz (1984) show, under the assumption of risk-neutrality, no payments should be transferred to the unemployed, as this would unnecessarily tighten the no-shirking condition. As a result, one only wants to maximise the utility of the employed workers in the steady state equilibrium. Taking the information and budget constraints into account, the social planner's maximisation problem can thus be written as

$$\begin{aligned} \max_{y,L,h} (y - c_E(h))L \quad \text{subject to} \\ y - c_E(h) \geq \frac{p+r+q(1-p)}{(1-p)(1-q)} \frac{C(h)}{D(h)} \\ \text{and} \quad yL \leq F(hL) \end{aligned} \quad (65)$$

Note that the probability of getting a job, p , is an increasing function of the employment level L . Using this, it follows that $\frac{p+r+q(1-p)}{(1-p)(1-q)} \frac{C(h)}{D(h)}$ is an increasing function of L . Let h_{SB} and L_{SB} denote the solutions to the above problem, i.e. the second-best optimal solutions of h and L respectively given the information constraint. It is easily verified that the two constraints in the above problem must be binding in optimum. From this it follows that

$$\frac{F(h_{SB}L_{SB})}{L_{SB}} = c_E(h_{SB}) + \frac{q+r+(1-q)p(L_{SB})}{(1-p(L_{SB}))(1-q)} \frac{C(h_{SB})}{D(h_{SB})}, \quad (66)$$

meaning that the average output per worker intersects the no-shirking condition. This is to be contrasted with the market outcome, where the marginal productivity per worker intersects the no-shirking condition, i.e.

$$h^*F'(h^*L^*) = c_E(h^*) + \frac{q+r+(1-q)p(L^*)}{(1-p(L^*))(1-q)} \frac{D(h^*)}{C(h^*)} \quad (67)$$

Due to the decreasing marginal productivity of work input we know that

$$hF'(hL) < \frac{F(hL)}{L} \quad \forall h > 0, \quad (68)$$

and from this it immediately becomes clear that we cannot have both $h_{SB} = h^*$ and $L_{SB} = L^*$. This holds regardless of our assumptions pertaining to the dependence of $\frac{C(h)}{D(h)}$ on h . Assume now the latter term does not depend on h . Then the optimal choice of L and h satisfies

$$F'(h_{SB}L_{SB}) = c'_E(h_{SB}), \quad (69)$$

i.e. the marginal productivity of each employee's work hours should equal his marginal disutility of working longer hours. A similar condition was derived in the market equilibrium,

which we repeat under the new assumption that $\frac{d}{dh} \left[\frac{C(h)}{D(h)} \right] = 0$:

$$F'(h^* L^*) = c'_E(h^*) \quad (70)$$

From these two equations we obtain the following result:

Lemma 2: *Under the above assumptions, the following holds:*

$$L_{SB} \leq L^* \Leftrightarrow h_{SB} \geq h^* \quad (71)$$

Proof: Assume $L_{SB} \geq L^*$ and $h_{SB} \geq h^*$ with strict inequality for at least one of them. Using the decreasing marginal productivity of work input it follows that

$$c'_E(h^*) = F'(h^* L^*) > F'(h_{SB} L_{SB}) = c'_E(h_{SB}), \quad (72)$$

which contradicts c'_E being increasing. Similar calculations can be performed in the case where $L_{SB} \leq L^*$ and $h_{SB} \leq h^*$, with at least one strict inequality. Hence the result follows. \square

We are now ready to prove the above Proposition:

Proof of Proposition 11: Assume that $L_{SB} < L^*$. Using that the no-shirking condition is binding, and the fact that $\frac{q+r+(1-q)p(L)}{(1-p(L))(1-q)}$ is an increasing function of L , it would thus follow that

$$\begin{aligned} (y_{SB} - c_E(h_{SB}))L_{SB} &= \frac{q+r+(1-q)p(L_{SB})}{(1-p(L_{SB}))(1-q)} L_{SB} \\ &< \frac{q+r+(1-q)p(L^*)}{(1-p(L^*))(1-q)} L^* = (y^* - c_E(h^*))L^* \end{aligned} \quad (73)$$

which would imply that the market outcome would dominate the solution to the social planner's problem. Since this cannot be the case, we must have a contradiction. We must thus have $L_{SB} \geq L^*$, and from the above lemma it now follows that $h_{SB} < h^*$, and the proof of the above Proposition is complete. \square

It follows that under the assumption that the derivative of $\frac{C(h)}{D(h)}$ is zero, the market choice of optimal individual work hours exceeds the second-best choice taking into account the information costs. It should be noted that, as in Shapiro and Stiglitz (1984), the second-best solution can be implemented by an appropriate tax on firms' profits complemented with a wage subsidy. Under the above assumptions, this would induce firms

to boost employment and simultaneously cut the length of the work-day.

In this section we have analysed the determination of individual working hours in the market in a simple shirking-model where employers' inability to perfectly monitor the workers' effort level leads to a sub-optimal aggregate level of employment. We found that, under certain conditions, the market outcome yields inefficiently long work day as compared to the first-best Pareto optimal level. However, this conclusion is also a possibility when we compare the market outcome to the second-best solution taking the market information costs into account: The market outcome could be improved upon, which could very well imply a decrease in individual work hours. To my knowledge, in their present form, neither of these results have been pointed out in the existing literature (but see Section 5.2).

5 Maximum hours regulation as policy response

The three previous sections have discussed and expanded some of the existing literature on individual overemployment. Even though some references have been made to the relevant empirical evidence, the main focus has been on theoretical analysis and modelling. In this section we comment on further empirical evidence which has quite frequently been cited in the literature. We also examine the suitability of a policy instrument which is often called upon by various commentators as a means to reduce the potential inefficiency associated with individual overemployment, namely the restriction of work hours through maximal hours legislation. As has already been noted in the literature on several occasions, such a policy may entail adverse effects which outweigh its potential benefits, a result that we also confirm.

5.1 Mismatch between the equilibrium allocation and the labour supply curve

As part of the analysis of the market determination of worktime length and its implications for various policy decisions, researchers have turned to the investigation of some of the empirical evidence on the labour market outcome. A notable example are the findings in Stewart and Swaffield (1997). They use data from the British Household Panel Survey for

1991 and find that 36% of British male employees aged 21-64 work longer hours than they prefer at the prevailing hourly wage rate. This is taken as evidence that wage packages offered by firms are bundled with minimum hours constraints. The authors further note:

The minimum hours constraints set by firms are an increasing function of the unemployment rate the individual faces. We hypothesise that this results from the individual's increased job insecurity and fear of redundancy and reduced alternative job opportunities at higher unemployment rates.

Moreover, the authors point out that these findings are at odds with the logic of the "canonical model of labour supply" of the Arrow-Debreu general equilibrium model, in which individuals choose the utility-maximising level of work hours for a given wage rate. They therefore dispute the relevance of the conclusions drawn from policy analysis based on this latter model.

Naylor (2003) proposes a framework within which one can highlight the difference between the Arrow-Debreu equilibrium and the description of the labour market suggested by the above evidence. Specifically, under the standard assumptions used in consumer theory, she describes an agent's utility-maximising choice of work hours for the given wage rate as the locus of turning points of the agent's indifference curves in the (h, w) -space, where h denotes work hours and w the hourly wage rate. This locus is referred to as the labour supply curve. Including the curve of the marginal product of labour (which is non-increasing in the (h, w) -space), the Arrow-Debreu equilibrium is described as the intersection of the latter curve with the agent's labour supply curve.

Naylor next introduces the idea of a *contract curve* which is defined as the locus of tangency points between the firm's isoprofit curves and the agent's indifference curves. From the point of the Arrow-Debreu equilibrium, this locus extends to the south-east in the (h, w) space. The locus represents the wage-hour bundles offered by a firm in a take it or leave it bargaining game between the firm and the agent under full information. Allocations on the contract curve below the Arrow-Debreu equilibrium point will thus only be attained in the case where the employer enjoys a degree of bargaining power and the agent's outside option is sufficiently low. One of the main points in Naylor (2003) is that the contract curve will lie to the south-east of the agent's labour supply curve. Hence in the case where the employer is able to push the agent off his labour supply curve, the agent will necessarily find himself in a situation where he would like to work fewer hours at the

wage rate associated with his current contract with the firm.

The reason for including a discussion of this empirical evidence and its treatment in the literature is to examine its relevance for economic policy analysis. In this respect it is worth noting some of the points which researchers have emphasised: For instance, Golden (1996) chooses to define the term *overemployment* as "workers reporting working in excess of their desired weekly or annual hours who are willing but unable to exchange (future) income for fewer hours within the firm or on the external labor market". One might contend that this choice of terminology has a normative touch to it. Still, the definition is largely consistent with the ideas discussed in the above references. Stewart and Swaffield (1997) further argue that "cutting hours may increase utility for many workers", whereas Naylor (2003) finds that maximum hours legislation will only raise workers' utility if it is combined with minimum wage regulations. Furthermore, both Rebitzer and Taylor (1995) and Landers et al. (1996) point out that the agents in their models "will perceive themselves to be hours constrained", although the relevance of the latter statement for their analysis is rather unclear.

We stated in the Introduction that our focus throughout the thesis would be on Pareto efficiency. Many of the above references do not mention this criterion in their respective discussions. Two comments are therefore warranted:

- First, even though maximum hours legislation and minimum wage regulation could increase workers' utility as Naylor (2003) argues, it should be noted that such policies could also potentially cause distortions in the market and result in the generation of an economic inefficiency in the Pareto sense, where individuals end up working too little. Consider for instance a standard take it or leave it bargaining model between a principal and an agent, much like the one we examined in Section 3.1. The principal owns a firm with production function $F(h)$, with $F'(h) > 0$, $F'' \leq 0$, where h denotes the agent's hourly input into production. The agent's utility is separable in the compensation, y , received from the principal, and his work input. We allow for the agent to be risk-averse. Thus, his utility function can be written as

$$u(y, h) = v(y) - c(h), \quad (74)$$

where the functions c and v satisfy $v(0) = c(0) = 0$, $v'(\cdot), c'(\cdot), c''(\cdot) > 0$ and $v''(\cdot) \leq 0$. To ensure the existence of an interior solution to the subsequent maximisation problems, we also assume that $\lim_{x \rightarrow 0+} xv'(x) > \lim_{y \rightarrow 0+} c'(y)$. Let w denote the hourly wage rate.

Under the above assumptions, the agent's labour supply curve in the (h, w) -space can easily be derived: The labour supply curve is found as the locus of points maximising utility subject to the agent's budget constraint for different values of the wage rate. That is, the agent's maximisation problem can be written as

$$\max_{h,y} v(y) - c(h) \quad \text{subject to } y = wh \quad (75)$$

Inserting the budget constraint into the utility function it is easily seen that the first-order condition associated with the above maximisation problem will satisfy $wv'(wh) = c'(h)$. Differentiating the first-order condition it is clear that

$$\frac{dw}{dh} = \frac{c''(h) - w^2 v''(wh)}{1 + hv''(wh)}, \quad (76)$$

as long as $1 + hv''(wh) \neq 0$. If the latter term is positive this indeed implies that the labour supply curve extends to the north-east. We will make this assumption to follow the discussion in Naylor (2003).

Now consider the principal's profit maximisation problem. We examine the case of perfect information where the agent has an outside option yielding a utility level $m \geq 0$. Hence the agent will never accept a contract from the principal if the implied utility level is lower than that of his reservation utility. The principal's problem can therefore be written as

$$\max_{h,y} F(h) - y \quad \text{subject to } v(y) - c(h) \geq m \quad (77)$$

Clearly, the constraint must be binding. Let h^* and y^* denote the solutions to the principal's problem. These then satisfy the following two conditions:

$$F'(h^*) = \frac{c'(h^*)}{v'(y^*)} \quad (78)$$

$$v(y^*) = m + c(h^*) \quad (79)$$

The implied optimal wage rate, $w^* := \frac{y^*}{h^*}$, is thus given by

$$w^* = \frac{v^{-1}[m + c(h^*)]}{h^*} \quad (80)$$

It should be noted that, except in the case of a risk neutral agent, the optimal level of work hours is generally a function of the agent's outside option, m . Keeping in mind that the second-order condition of the principal's maximisation problem requires

$$F''(h) - \frac{c''(h)}{v'(y)^2} - \frac{c'(h)^2 v''(y)}{v'(y)^3} < 0, \quad (81)$$

it follows that an increase in the agent's reservation utility level m implies a reduction in the optimal number of hours:

$$\begin{aligned}\frac{dh^*}{dm} &= -\frac{\frac{c'(h^*)^2 v''(y^*)}{v'(y^*)^3}}{F''(h^*) - \frac{c''(h^*)}{v'(y^*)} - \frac{c'(h^*) v''(y^*)}{v'(y^*)^3}} \\ &= -\frac{c'(h^*) v''(y^*)}{v'(y^*)^3 F''(h^*) - v'(y^*)^2 c''(h^*) - c'(h^*)^2 v''(y^*)} < 0\end{aligned}\tag{82}$$

Thus, under the present assumptions of separable utility, the contract curve as described by Naylor (2003) would be extending in an eastward direction in the (h, w) -space. Now on the one hand, as long as the agent's outside option yields a utility level so low that h^* and y^* satisfy

$$\frac{c'(h^*)}{v'(y^*)} > \frac{y^*}{h^*} = w^*,\tag{83}$$

the contract curve will lie to the right of the labour supply curve and the agent would prefer to work fewer hours at the wage rate implied by the contract offered by the principal. On the other hand, under the present assumptions of perfect information and the absence of externalities, the contract curve corresponds exactly to the locus of all Pareto efficient allocations in the economy. Three points must be emphasised in this context:

First, as was pointed out in Naylor (2003), legislation restricting the number of hours will not alone be sufficient to raise the welfare of the worker. Consider the case where working hours are required not to exceed an upper level \bar{h} . When this constraint is binding, the principal's profits are simply given by $F(\bar{h}) - w/\bar{h}$. Suppose the Agent's participation constraint is not binding. Then a simple reduction in the hourly wage rate such that the participation constraint is still satisfied would lead to an increase in the principal's profits. It follows that the only sub-game perfect Nash equilibrium leaves the agent at his reservation utility level m . Thus, increasing the utility of workers may require the combination of maximum hours legislation *together* with minimum wage regulation.

Second, as Naylor also briefly mentions, introducing maximum hours legislation is likely to cause the bargaining outcome between the principal and the agent to occur off the original contract curve. What Naylor fails to point out is that this involves moving the economy off the locus of Pareto efficient equilibria. We have the following result:

Proposition 12: *With the absence of externalities and under perfect information, the introduction of maximum hours legislation as the sole policy intervention will create an*

inefficient market outcome in which the agent works too little, without increasing the latter's utility.

Proof: We assume that the hours restrictions are binding, meaning $\bar{h} < h^*$. As we already argued the agent's participation constrain must be binding after the policy intervention. Let \tilde{y} denote the new payment to the agent. Since c is increasing, we must thus have

$$v(\tilde{y}) = m + c(\bar{h}) < m + c(h^*) = v(y^*), \quad (84)$$

which implies $\tilde{y} < y$. As a consequence, since c' is increasing and both v' and F' are decreasing, it follows that

$$F'(\bar{h}) > F'(h^*) = \frac{c'(h^*)}{v'(y^*)} > \frac{c'(\bar{h})}{v'(\tilde{y})} \quad (85)$$

This means that the value of the agent's marginal product exceeds the amount the agent would need to be willing to increase his labour supply, meaning that the agent indeed works too little. \square

It follows that maximum hours legislation as a sole policy intervention will not in general be a first-best response if the aim is to increase workers' welfare. It may even result in economic inefficiencies, without reaching its aims of increased worker welfare. As we have already seen, in this framework of perfect information that we are considering, instead of directly regulating the length of the work-week a policy that focuses on improving the agents' outside option, that is, increasing m , will automatically yield a market outcome with a reduction of work hours along the locus of Pareto efficient equilibria.

Third, and lastly, it should now be clear that employees' reporting of working in excess of their desired hours at the current wage rate should not necessarily be interpreted as evidence of an economic problem, at least not in the Pareto sense. To the extent that such evidence calls for a legislative intervention, one might argue that the appropriate conduct would be to find a politically desirable point on the contract curve, and not enforcing regulations which moves the economy away from this locus of Pareto efficient equilibria. The idea is summarised below:

Proposition 13: *Employees reporting to work beyond their desired number of hours at the prevailing wage rate is not in itself evidence of inefficiency due to individual overemployment: Even when the employers' bargaining power enables him to push his employees off their labour supply curve, the resulting allocation may still be on the economy's locus*

of Pareto efficient equilibria. It follows that employers' market power need not in itself be a source of economic inefficiency regarding the determination of individual work hours.

It should be emphasised that the validity of the above two Propositions is less obvious in a setting characterised by externalities and/or imperfect information. Our second comment on the evidence presented in Stewart and Swaffield (1997), however, is valid also under more general conditions, and is also more straightforward:

- Since the authors focus on the fact that some individuals prefer to work fewer hours at their current wage rate, one might falsely get the impression that the concern for overemployment as we have defined it in this thesis is of less relevance in the absence of such evidence. However, from the analyses in the previous sections it should be clear that this is not the case. Indeed, in several of the models that we have examined inefficiency due to individual overemployment occurs despite the fact that individuals would not be willing to work fewer hours at the prevalent wage rate. This is particularly the case with the conspicuous consumption model of Section 2 and the career concerns models of Section 3. Our analysis therefore suggests that such reports from workers are of minor importance as empirical evidence of inefficiency associated with individual overemployment.

The above examination of the empirical evidence presented in Stewart and Swaffield (1997) may be misinterpreted as doubting the general relevance of such evidence. However, this is clearly not the case. As Golden (1996), Stewart and Swaffield (1997), and Naylor (2003) all point out, the finding that individuals do not find themselves on their labour supply curve should be thought of as convincing evidence that the canonical Arrow-Debreu model of labour supply is a poor description of reality. This further has the implication that one is forced to question the adequacy of a straightforward normative conclusions obtained from policy analysis based on this latter model.

5.2 Work-sharing as policy instrument against overemployment

During our analysis of the shirking model presented in Section 4.2.2 we noted that, under the condition of decreasing marginal productivity of effective labour, the market outcome was not unlikely to yield a level of individual work hours which exceeds both the first-best Pareto efficient level as well as the second-best level where the information costs were taken into account. We pointed out that a policy that boosted employment levels could potentially bring the market outcome closer to both the first- and the second-best optima

with respect to individual work hours. Similar ideas have been discussed at length in the literature in a variety of different settings. However, in many cases, the focus has been on the reverse direction of causality, that is, the extent to which a reduction in individual work hours can actively be used as an instrument to decrease unemployment, an idea which is often referred to as *work-sharing*. And assuredly, many have shared the conviction that work-sharing is both a suitable and desirable measure to fight unemployment, an idea which is perhaps best illustrated by the founder of the *American Federation of Labor* Samuel Gompers's asseveration in 1887 that

So long as there is one man who seeks employment and cannot find it, the hours of work are too long.

(As quoted in Earle and Pencavel (1990))

In addition to its assumed positive effect for the individual worker, many have thus argued that maximum hours legislation should be used to combat unemployment. The conclusions drawn in the economics literature from the analysis of this proposal have been rather ambiguous, however. Although several authors have found cases where regulations limiting the level of standard work hours may indeed result in a reduction in unemployment, many have also found that such regulations might have the complete opposite effect: Regulating the maximum number of work hours may increase firms' costs and hence reduce their demand for labour. At an aggregate level, this market intervention may thus cause an increase in the unemployment rate. See Hoel and Vale (1986), Hoel (1987), Hunt and Katz (1998), and Freeman (1998) for some examples from this literature.

It should come as no surprise that the effect of introducing maximum hours legislation on the unemployment rate depends crucially on the underlying reason for the very existence of unemployment in the first place. Our shirking-model of Section 4.2.2 may serve to illustrate the potential adverse effects on the employment levels of such a policy. Recall that the steady state free-market equilibrium level of individual work hours and employment were denoted by h^* and L^* respectively. Consider introducing a legislation restricting work hours not to exceed a level $\bar{h} < h^*$, such that this new constraint is binding. It should be clear that the no-shirking condition in (57) will also still be binding. With the assumption that the right-hand side of (60) is positive for all h , that is

$$c'_E(h) + \frac{p+r+q(1-p)}{(1-p)(1-q)} \frac{C(h)}{D(h)} \left[\frac{C'(h)}{C(h)} - \frac{D'(h)}{D(h)} \right] > 0 \quad \forall h > 0, \quad (86)$$

it should be clear that introducing this legislation alone will lead to a decrease the compensation paid to the individual worker, given the initial unemployment rate. This effect would in isolation provide the firm with incentives to increase its demand for labour. However, due to the hours restriction, the gain from hiring more workers may also be smaller. Whether aggregate employment increases or not depends on which effect is dominating, and is summarised in the following result which needs no proof:

Proposition 14: *Let \tilde{L} denote the firms' optimal demand for labour after the introduction of maximum hours legislation with $\bar{h} < h^*$. Then*

$$\tilde{L} \leq L^* \Leftrightarrow \bar{h}F'(\bar{h}L^*) \leq c_E(\bar{h}) + \frac{r + q + (1 - q)p(L^*)}{(1 - p(L^*))(1 - q)} \frac{C(\bar{h})}{D(\bar{h})}, \quad (87)$$

that is, aggregate employment increases if and only if the reduction in worker remuneration offsets the potential decrease in labour productivity.

In our framework the consequences for the individual firm's labour demand (and thus the aggregate unemployment rate) depend crucially on the functional form of the term $\frac{c_E(h) - c_S(h)}{D(h)}$ (which determines the size of the change in the workers' compensation). The main point here is that a decrease in the unemployment rate cannot be guaranteed after introducing maximum hours legislation only. To what extent such a policy will result in a potential Pareto improvement is thus also not clear. As noted in Section 4.2.2, a policy subsidising the workers' compensation might be more suited to increase employment in this particular model, with the possible added benefit that individual work hours may also decrease.

6 Concluding remarks

The purpose of this thesis has been to present and discuss the existing frameworks within which individual overemployment can potentially occur. The main focus has been on theoretical models which allow for such a possibility. The most obvious circumstance under which such overemployment would occur is under the presence of externalities. Here, we have examined one particularly interesting case, in which the externality emanates from the agents' concern for their relative status within society. Furthermore, as we have seen, the classical condition of asymmetric information which traditional analyses have found will often lead to market failure characterised by underemployment can also easily

be used to produce equilibria where individuals work too much. In some cases the market mechanisms may cause all affected individuals to work hours or exert effort beyond any level required for Pareto efficiency. Often, however, the market outcome will simultaneously exhibit both underemployment and overemployment.

In the introduction we stated that one of the motivations for investigating the conditions under which Pareto inefficiencies associated with individual overemployment occur was to examine the extent to which *exploitation* and *exhaustion* of labour is likely to involve such inefficiencies. It is natural in this context to discuss which segments of the labour market are most likely to be characterised by individual overemployment and whether the efficiency merits of labour markets are likely to change during the course of economic development and growth. It is worth noting that many of the models analysed in this thesis seem to be constructed with a view to describing overemployment in the upper "strata" of society. The career concerns models discussed in Section 3 seem to fit this description well. For instance, the original model by Holmström (1982) sought to analyse the actions and compensation of firms' managers, whereas the model by Kaarbøe and Olsen (2006) was developed to determine the optimal design of monetary incentives for medical doctors working in hospitals. Landers et al. (1996) analyse the determination of work hours in law firms. Rebitzer and Taylor (1995) note that a plausible interpretation of their model is that it "[only applies] to the high-paying 'primary' sector jobs." Moreover, to the extent that one can assume that an individual's concern for his relative position in society only come into play once the individual has already covered his basic essential needs and reached a state beyond subsistence, one might argue that the model of conspicuous consumption analysed in Section 2 perhaps is more relevant in prosperous industrialised rather than poor developing countries. Furthermore, in the original rat race model by Akerlof (1976) only talented workers with high pay end up working beyond the Pareto efficient level.

The existing academic treatment and theoretical analysis in economics of the notion of overemployment thus seems to support its potential existence in both high-paying sector jobs in particular, and in rather rich countries in general. On the other hand, much of the discussion regarding overemployment in media seems to pertain to the case of low-paying sector jobs, especially those found in relatively poor developing countries. In particular, the allegation by certain commentators that the latter jobs represent both exploitation

and exhaustion of labour does not seem to fit as well into the various models presented and discussed thus far. One might therefore question the extent to which situations as described in for instance Amnesty International (2013) do in any way constitute inefficiencies due to individual overemployment in the Pareto sense. It may very well be that other criteria, possibly originating from a quite different ethical framework, are needed to successfully establish academically valid arguments capable of denouncing the current labour market situation found in many developing countries.

References

- Akerlof, George (1976). "The economics of caste and of the rat race and other woeful tales". *Quarterly Journal of Economics* 90(4), 599-617
- Amnesty International (2013). *The Dark Side of Migration: Spotlight on Qatar's Construction Sector Ahead of the World Cup*. London: Amnesty International Ltd.
- Arrow, Kenneth J., and Partha S. Dasgupta (2009). "Conspicuous Consumption, Inconspicuous Leisure". *The Economic Journal* 119.541: F497-F516.
- Banerjee, Abhijit V., Paul J. Gertler, and Maitreesh Ghatak (2002). "Empowerment and efficiency: tenancy reform in West Bengal". *Journal of political economy* 110.2: 239-280.
- Bardhan, Pranab, and Christopher Udry (1999). *Development microeconomics*, Oxford University Press.
- Bolton, Patrick and Mathias Dewatripont (2005). *Contract Theory*. Cambridge, Massachusetts: The MIT Press.
- Boskin, Michael J. and Eytan Sheshinski (1978). "Optimal redistributive taxation when individual welfare depends upon relative income". *Quarterly Journal of Economics* 92(4), 589-601.
- Bowles, Samuel (1981). "Competitive wage determination and involuntary unemployment: A conflict model". mimeo. University of Massachusetts.

- Burdett, Kenneth and Dale T. Mortensen (1980). "Search, layoffs, and labor market equilibrium". *Journal of Political Economy* 88(4), 652-672.
- Cahuc, Pierre and Fabien Postel-Vinay (2005). "Social status and the overworked consumer". *Annals of Economics and Statistics* (78), 143-161.
- Clark, Andrew E., Paul Frijters, and Michael A. Shields (2008). "Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles". *Journal of Economic Literature*: 95-144.
- Dewatripont, Mathias, Ian Jewitt, and Jean Tirole (1999a). "The Economics of Career Concerns, Part I: Comparing Information Structures". *Review of Economic Studies*, 66: 183-198
- Dewatripont, Mathias, Ian Jewitt, and Jean Tirole (1999b). "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies". *Review of Economic Studies*, 66: 199-217
- Diamond, Peter A. and Eric Maskin (1979). "An equilibrium analysis of search and breach of contract, i: Steady states". *Bell Journal of Economics* 10(1), 282-316.
- Diamond, Peter A. (1982). "Wage Determination and Efficiency in Search Equilibrium". *Review of Economic Studies* 49(2), 217-227
- Duesenberry, James S. (1949). *Income, Saving and the Theory of Consumer Behavior*, Harvard: Cambridge
- Earle, John S. and John Pencavel (1990). "Hours of work and trade unionism". *Journal of Labor Economics* 8(1), S150-174.
- Easterlin, Richard (1974). "Does economic growth improve the human lot? Some empirical evidence". In P. A. David and M. W. Reder (Eds.), *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*, Volume 89. New York: Academic Press.
- Easterlin, Richard A., Laura A. McVey, Malgorzata Switek, Onnicha Sawangfa, and Jacqueline S. Zweig. "The happiness-income paradox revisited". *Proceedings of the National Academy of Sciences* 107, no. 52 (2010): 22463-22468.

- Fisher, Walter H., and Franz X. Hof (2000). "Relative consumption, economic growth, and taxation". *Journal of Economics* 72.3: 241-262.
- Frank, Robert H. (1991). "Positional Externalities", in Richard Zeckhauser, ed., *Strategy and Choice: Essays in Honor of Thomas C. Schelling*, Cambridge, MA: MIT Press: 25-47.
- Freeman, Richard (1998). "Work-Sharing to Full Employment: Serious Option or Populist Fallacy?" In *Generating Jobs: How to Increase Demand for Less-Skilled Workers*, edited by Richard Freeman and Peter Gottschalk. New York: Russell Sage Foundation.
- Gibbons, Robert and Kevin J. Murphy (1992). "Optimal incentive contracts in the presence of career concerns: Theory and evidence". *Journal of Political Economy* 100(3), 468-505.
- Golden, Lonnie (1996). "The economics of worktime length, adjustment, and flexibility: A synthesis of contributions from competing models of the labor market". *Review of Social Economy* 54(1), 1-45.
- Hagerty, Michael R., and Ruut Veenhoven (2003). "Wealth and happiness revisited - growing national income does go with greater happiness". *Social indicators research* 64.1: 1-27.
- Hansen, Gary D. (1985). "Indivisible labor and the business cycle". *Journal of Monetary Economics* 16, 309-327.
- Hoel, Michael and Bent Vale (1986). "Effects on unemployment of reduced working time in an economy where firms set wages". *European Economic Review*, 30(5), 1097-1104.
- Hoel, Michael (1987). "Can Shorter Working Time Reduce Unemployment?" In *Unemployment in Europe: Analysis and Policy Issues*, edited by Claes Henric Siven. Stockholm: Timbro.
- Holmström, Bengt (1982). "Managerial Incentive Problems - A Dynamic Perspective". In *Essays in Economics and Management in Honor of Lars Wahlbeck*, Helsinki: Swedish School of Economics.

- Hunt, Jennifer and Lawrence F. Katz (1998) "Hours reductions as work-sharing". *Brookings papers on economic activity* 339-381.
- International Labor Organization (2013). *Case Study: Karoshi: Death from overwork*, http://www.ilo.org/safework/info/publications/WCMS_211571/lang-en/index.htm (accessed 14/5-2014)
- Kaarbøe, Oddvar M. and Trond E. Olsen (2006). "Career concerns, monetary incentives and job design". *The Scandinavian Journal of Economics* 108 (2), 299-316.
- Laffont, Jean-Jacques and David Martimort (2002): *The Theory of Incentives: The Principal-Agent Model*, Princeton University Press
- Landers, Renée M., James B. Rebitzer, and Lowell J. Taylor (1996). "Rat race redux: Adverse selection in the determination of work hours in law firms". *American Economic Review* 86(3), 329-348.
- Ljungqvist, Lars, and Harald Uhlig (2000). "Tax policy and aggregate demand management under catching up with the Joneses". *American Economic Review* 90.3: 356-366.
- Lucas, Robert E. and Edward C. Prescott (1974). "Equilibrium search and unemployment". *Journal of Economic Theory* 7(2), 188-209.
- Luttmer, Erzo F.P. (2005) "Neighbors as negatives: Relative earnings and well-being". *The Quarterly Journal of Economics* 120.3: 963-1002.
- Mortensen, Dale T. (1970). "Job search, the duration of unemployment, and the phillips curve". *American Economic Review* 60 (5), 847-862.
- Naylor, Robin A. (2003). "Labour supply, efficient bargains and countervailing power". In M. Waterson (Ed.), *Competition, monopoly and corporate governance: Essays in honour of Keith Cowling*, pp. 100-116. Edward Elgar Pub.
- Pauly, Mark V. (1968) "The economics of moral hazard: Comment". *The American Economic Review*: 531-537.
- Pissarides, Christopher A. (1979). "Job matchings with state employment agencies and random search". *The Economic Journal* 89 (356), 818-833.

- Rebitzer, James B. and Lowell J. Taylor (1995). "Do labor markets provide enough short-hour jobs? An analysis of work hours and work incentives". *Economic Inquiry* 33(2), 257-273.
- Robson, Arthur J. (1992) "Status, the distribution of wealth, private and social attitudes to risk". *Econometrica: Journal of the Econometric Society*: 837-857.
- Rogerson, Richard (1988). "Indivisible labor, lotteries and equilibrium". *Journal of Monetary Economics* 21(1), 3-16.
- Schor, Juliet (1993). *The overworked American: The unexpected decline of leisure*. Basic books.
- Shapiro, Carl and Joseph E. Stiglitz (1984). "Equilibrium unemployment as a worker discipline device". *American Economic Review* 74 (3), 433-444.
- Stewart, Mark B. and Joanna K. Swaffield (1997). "Constraints on the desired hours of work of British men". *The Economic Journal* 107, 520-535.
- Stoft, Steven (1982). "Cheat-threat theory: An explanation of involuntary unemployment". mimeo. Boston University.
- Veblen, Thorstein (1899) *The Theory of the Leisure Class: An Economic Study of Institutions*, Macmillan
- Veenhoven, Ruut, and Michael Hagerty (2006). "Rising happiness in nations 1946-2004: A reply to Easterlin". *Social Indicators Research* 79.3: 421-436.
- Weiss, Yoram, and Chaim Fershtman (1998). "Social status and economic performance: A survey". *European Economic Review* 42.3-5: 801-820.
- Weitzman, Martin L. (1982). "Increasing returns and the foundations of unemployment theory". *The Economic Journal* 92, 787-804.
- Yellen, Janet L. (1984). "Efficiency wage models of unemployment". *American Economic Review* 74(2), 200-205.